



Thomas, P. J. (2020). Minimum sample size for the survey measurement of a wealth-dependent parameter with the UK VPF as exemplar. *Measurement*, 150, [107044].
<https://doi.org/10.1016/j.measurement.2019.107044>

Publisher's PDF, also known as Version of record

License (if available):
CC BY

Link to published version (if available):
[10.1016/j.measurement.2019.107044](https://doi.org/10.1016/j.measurement.2019.107044)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the final published version of the article (version of record). It first appeared online via Elsevier at <https://www.sciencedirect.com/science/article/pii/S0263224119309108> . Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>



Minimum sample size for the survey measurement of a wealth-dependent parameter with the UK VPF as exemplar

Philip Thomas

Risk Management, Safety Systems Research Centre, South West Nuclear Hub, Faculty of Engineering, University of Bristol, Tower View, Queen's Building, University Walk, Bristol BS8 1TR, United Kingdom

ARTICLE INFO

Article history:

Received 6 May 2019

Received in revised form 19 August 2019

Accepted 9 September 2019

Available online 12 September 2019

Keywords:

Measurement science

Economic measurement

Sample size

Opinion survey

Opinion poll

Wealth

Value of a prevented fatality

VPF

ABSTRACT

Measurement of an economic good by opinion survey constitutes a variant of the political opinion polls widely familiar from news reporting. The paper relates the minimum sample size needed for the survey measurement of a wealth-dependent parameter to the smallest sample for a political poll giving the same precision. Measuring a strongly wealth-dependent parameter by survey requires a sample size of ~2000 or more to provide precision equivalent to the 3% margin of error customary in UK political opinion polls. It is shown that the survey measurement of the “value of a prevented fatality” (VPF) used in the UK as a health and safety spending yardstick requires ~3000 people to be questioned. The analysis shows the actual sample size used, 167, to be inadequate. This adds to the problems besetting the UK VPF, as the method the surveyors used to interpret their data has already been shown invalid.

© 2019 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Measurement history in the social sciences

It is widely acknowledged that metrology has played an enormous role in the development of the physical sciences that have seen such success over the last 150 years. Less well appreciated is the fact that measurement has been applied to commerce since the earliest days of man's civilisation [1,2]. Furthermore, increasing attention is now being paid to the processes by which human attributes such as perceptions, judgements and valuations might be measured. There can be no doubt, however, that measuring the properties or attributes of human behaviour is a difficult task.

The economist Edward Glaeser argues [3] that the social sciences are still young, claiming:

“The widespread application of scientific methods to the study of human society—rigorous formal theories, serious empirical testing—occurred only during the twentieth century, mostly since World War II.”

He highlights the work of “great measurers” of the 20th century such as the economist Simon Kuznets, whose economic tradition he expects will continue to expand in the coming decades. Kuznets

stressed that reliable results can be derived only through large numbers of observations and his Nobel prize winning work in economics provided the key to measuring gross national product [4,5]. The need for large numbers of observations will be a theme to which this paper will return.

But measurement on a large scale in the field of social science dates back at least to the middle of the nineteenth century, when the first public examinations for schools were introduced in the United Kingdom. The move was in response to a demand for some way of judging pupils' levels of attainment [6] and in 1858, the University of Cambridge Local Examinations Syndicate set and marked papers for school students under the ages of 16 and 18 over a wide range of topics. The subjects were similar to those examined today: English Language and Literature, History, Geography, Geology, Greek, Latin, French, German, Political Economy and English Law, Mathematics, Arithmetic, Chemistry, Physical Sciences, Zoology, Drawing, Music and Religious Knowledge. The issues involved in measuring proficiency from the answers to graded questions have been the subject of a number recent research studies [7,8]. Meanwhile a good summary of the historical development of measurement thinking since the 1930s for more general psychological attributes is given in [9]. Deriving from the success of measurement in the physical sciences, a philosophical and theoretical basis was developed for the universal process of

E-mail address: philip.thomas@bristol.ac.uk

measurement [10]. A proposal for a formal theory of a general measurement system [11] has been put forward recently.

The need to determine what can and cannot be claimed for measurements in the social sciences has stimulated the re-examination of the principles of metrology. Stevens made a notable contribution just after the Second World War [12] by analysing scales of measurement and classifying them as nominal, ordinal, interval and ratio. The ratio scale is the most comprehensive of these, in the sense that it can accommodate the four important quantitative relations: equality, rank-order, equality of intervals and equality of ratios. At about the same time, von Neumann and Morgenstern [13] examined the principles of measurement in relation to the concept of utility widely used in economics then and now. They concluded that utility could be claimed correct to a positive linear transformation – Stevens's interval scale. Subsequently Thomas [14,15] was able to build on this work by including the notions concerning utility introduced by Pratt [16] and Atkinson [17] to show that a ratio scale for utility against wealth can be constructed, with risk-aversion as parameter.

Schley and Peters [18] have proposed that an individual's risk-aversion is influenced by the subject's imprecise perception of numerical magnitudes: "inexact mappings of symbolic numbers onto mental magnitudes". Differences in the curvatures of people's utility functions can then be explained, at least in part, by differing levels of numerical acuity. Those with higher powers of numerical discrimination display an effective risk-aversion that is lower than those with a less well developed ability to distinguish between numbers of a similar but different magnitude. Meanwhile, different situations call for different risk-aversions. Thomas and Chrystal [19] suggest that a consumer will be open to quantity retail promotions such as "buy one get one free" provided his/her risk-aversion lies between 0 and 1. When commodities such as eggs are sold in different pack sizes, satisfying the desires of the average cautious consumer (risk-aversion = 0.5) will then result in a ratio of successive pack sizes equal to the square of the golden ratio, namely 2.62, while the price-ratio will be the golden ratio, 1.62. The "three for the price of two" promotion (where, for example, 18 eggs are offered for the price of 2 half-dozen packs) emerges as a reasonably close approximation, with an implied risk-aversion of 0.37.

Thomas [15] laid out the challenge of measuring risk-aversion, an important parameter lying squarely in the field of social science, at the interface between psychology and economics. Further progress was made when it was shown recently that this psycho-economic property can be measured empirically for decisions on life extension using data collected from the lives of billions of people throughout the world [20]. See also [21]. The measurement became possible as the result of the application of a new theory, the J-value [22,23], to decisions on extending life. The study by Thomas and Waddington utilised the results of the continuing "natural experiment" that is taking place in all societies at all times to find the optimal trade-off between promoting longer life (for example through better public health, enhanced health care and improved industrial protection) and the resources that people are prepared to devote to the task, under the constraint that different nations have different levels of resource.

1.2. Problems facing measurement by survey

Subjectivity and bias are two obvious hazards lying in wait for those attempting measurements in non-physical fields. Finkelstein [2] suggests that the observers/analysts in social sciences may well not be objective, but operate "on the basis of ideologically motivated theories". Mari et al. [24] follow Karl Popper [25] in regarding objectivity and inter-subjective testability as the critical features for the reliability and dependability of physical and

non-physical measurements. Transgressions against those precepts can arise in one particular method of measurement used extensively in the social sciences, namely an opinion survey, by which the consolidated view of the wider population is sought by questioning only a sample. The intention is usually to measure a subjective parameter, and this may bring its own problems, as discussed later.

The important concept of the "person as a measuring instrument" is introduced by Rossi and Berglund [9] in the context of human perceptions of smell, who propose that "the screening and testing of participants, as measuring instruments, are prerequisites for reliable and valid psychological measurement". An example would be the wine tasters employed in wineries and restaurants, who, by nature and experience, have developed more sophisticated and discerning palates than the average person and are therefore able to pass a more expert verdict. Their judgements will influence the market into which a particular wine is sold, as well as its price. The wine-tasting process might be automated eventually, perhaps using a version of the electronic nose discussed in [9], possibly acting in conjunction with analytic chemical instrumentation.

After "sentencing" by the experts, the wine will subsequently be tasted by many people, and each will make his or her own evaluation of it. It is important to note that, in this case, each person will be making a measurement of the same thing, namely the qualities of the wine in question. However, this contrasts strongly with many survey measurements made in the socio-political arena, where the person is certainly the measuring instrument, but what is being measured is not common to all but is, instead, unique to the person.

For example, by the theory of the utility model for the value of a prevented fatality (VPF) (see Section 3.1 for details), the individual is required to possess a maximum acceptable price (MAP) that he or she is prepared to pay to avert an injury and a minimum acceptable compensation level (MAC) that he/she is prepared to accept to make up for receiving that injury. It should be noted that the MAP and the MAC are specific to the individual and cannot be regarded as the individual's subjective measurement of a feature which the other respondents have access to and can measure. In these surveys, there is no better expert on the opinion of the individual than the individual himself/herself. Moreover, screening out individuals in such a case is not justified unless it can be proved in advance that the person is not qualified to offer an opinion, for example by reason of insanity or through being too young to have acquired the necessary experience (e.g. being below voting age). There is no correct view in such a survey, and in an open and civilised society, there should be no discrimination in favour of certain, selected opinions. It is with such surveys that the present paper is concerned.

The least problematic and simplest of such surveys is probably the political opinion poll conducted among adults qualified to vote. This is because in a democracy, the respondent has exclusive control of his or her opinion, which he or she may later convert into a vote. But even here the process of conversion from opinion to vote is not straightforward. For example, the person's view can change between the date of the survey and the date of the election, and, of course, the person may not vote on the day, either through choice or circumstance. Moreover, the respondent may not want to divulge his/her true opinion to the pollster and may instead express a view calculated to avoid the interviewer's presumed censure.

It is obvious that, as a minimum, the subjective opinions of the surveyor should not be allowed to affect the result of the political opinion survey. There are two correctives against this. In the first instance, it is unlikely in a free and competitive society that there

will be only one political polling organisation, so that a fairly immediate check on results will be at hand. Moreover there will eventually be a definitive check when the result of the official ballot becomes known. But it is important to realise that such checks and balances may well be absent in surveys beyond the political arena, and this increases the requirement for the highest levels of impartiality among surveyors and analysts.

The need to avoid analysts' opinions encroaching on the results of their opinion surveys led to the development in 2014 of the new criterion of Structural View Independence [26]. The mathematics prove that it is possible for the analyst to select in advance a plausible method to consolidate the survey results into a single figure so that the final value is guaranteed to be biased either low or high. For example, the geometric mean will always produce a figure that is too low, while the root-mean-square value will always give a value that is too high. These are instances where a general non-linear, increasing and differentiable transformation is applied, then the mean found for the transformed sample and finally the result back-transformed to give the consolidated figure. Structural View Independence requires the study to be structured to ensure that the weighting given to each person's opinion is independent of its content. Only in this way can the views of all respondents be accorded the equal treatment they deserve. Out of all the transformations considered, this criterion is satisfied only by linear transformations. This implies that the consolidated view should be found simply by taking the arithmetic average. In one important case, a team of surveyors employed a further, non-differentiable transformation to interpret opinion surveys, and the results were used to justify reducing the spending against multi-fatality accidents on the UK's railways by a factor of three. However it was shown in [27] that this new transformation discriminated systematically in favour of low valuations (and hence low spendings) and thus violated the principle of Structural View Independence. This meant that it was unsafe to make use of the survey results.

It is obvious from the discussion above that there are significant pitfalls associated with the application of opinion surveys for general parameter estimation. Even so they are often used in economics to put a value on public goods, such as clean air or the continued survival of a rare species of plant or animal.

It is generally accepted [28] that the ways to measure the value of any good are, in order of preference:

1. the market value if a free market in the good exists – this is the best way
2. the value deduced from revealed preferences, in line with John Locke's dictum "I have always thought the actions of men the best interpreters of their thoughts" [29]
3. the value deduced from stated preferences, typically from opinion surveys – where Fujiwara and Campbell [30] have uttered the reservation: "respondents in stated preference surveys may have an incentive to deliberately misrepresent their true preferences in order to achieve a more desirable outcome for themselves ... individuals may overstate their valuations of the good if they believe their responses influence its provision and are unrelated to the price they will be charged for it".

It is when methods 1 and 2 prove either impossible or very difficult to arrange that option 3 is sometimes chosen.

As a minimum the criterion of Structural View Independence needs to be satisfied if an opinion survey is to be used as a measurement tool. It is equally clear that the most rigorous standards must be applied to the task of statistical inference if the survey results are to have meaning. Even then, of course, the Fujiwara and Campbell caveat suggests that that high accuracy cannot be expected.

1.3. Survey measurement of the "value of a prevented fatality"

The present author has highlighted a number of instances of poor practice in survey measurements, where lack of rigour has had notable adverse implications [31]. One ethically significant and practically important case concerns the measurement of the UK Government's "value of a prevented fatality" (VPF). This is defined as the maximum amount that it is notionally reasonable to pay for a safety measure that will reduce by one the expected number of preventable premature deaths in a large population. But it is pointed out in [32] that the VPF constitutes only a crude estimate of what is lost when someone is deprived of his or her life. For example, an 18 year-old generally has many more years of life ahead than a 78 year-old. Moreover, the UK VPF, which is based on an opinion survey carried out over 20 years ago [33], has been shown to contain major flaws as a result of the use of an invalid method of interpreting its survey evidence [34].

The success in validating the J-value method against pan-national data [20] enables the statement to be made that the UK VPF in current use, £1.83M (2016 £s), is 4 to 5 times below what would be reasonable. To the extent that the concept has validity, the VPF should lie a lot closer to the equivalent figure announced by the U.S. Department for Transportation in 2016, namely \$9.6M, based on stated preference studies [35]. Using J-value analysis for valuing human life, the VPF, now interpreted as the amount that should be spent to preserve the life expectancy of the average person in the UK, ought to be £8.59M (2015 £s). The clear disparity between this figure and that still used by the UK Government has obvious, negative implications for the priority being assigned to the safety of UK citizens.

Although the research team responsible for the VPF survey made two attempts to defend its interpretation method, [36,37], the team's defence was refuted in detail [38,39]. See also [40] for a review of the multiple flaws besetting the UK VPF and [41] for a demonstration of the lack of justification for the surveyors rejecting their initial survey, which would have produced a much higher figure [42].

It has been pointed previously that the validity of a survey estimate for the VPF "rests critically on the sample population reflecting closely the probability density for wealth of the target population as a whole" (last paragraph of Appendix A.5 of [34]). But the average wealth of the participants in the Carthy survey may be calculated using the surveyors' own method as between £3568 and £7136, less than 10% of the average net wealth of UK adults at the time of the survey.

A further factor contributing to the glaring mismatch between the apparently low levels of wealth amongst the respondents and the much higher figures in the UK population as a whole might well be the small sample size used in the VPF survey: just 167 people.

The important general question is raised: what is the minimum sample size needed to measure a wealth-dependent economic parameter by opinion survey? In addressing this issue, the well documented case of the survey measurement of the VPF is used as an exemplar. Such a survey is a variant of the political opinion polls widely familiar from news reporting. A significant difference is that the measurement of the economic good usually involves the characterisation of people's choices from a continuum rather than from two or more discrete options.

1.4. Structure of the paper

This section has provided a brief review of the history of measurement in the social sciences, of the difficulties encountered in using surveys as measurement tools and of some the problems that

have been observed in attempts to measure the “value of a prevented fatality” (VPF).

New data on the distribution of personal wealth in the UK allow an assessment to be made, in Section 2, of the challenge to the measurement of a wealth-dependent parameter that is posed by the very great variation in wealth across a developed country such as the United Kingdom. Section 3 develops two diverse models for the survey measurement of the VPF: the Utility Model and the Multiplier Model.

In Section 4, the smallest sample size needed to guarantee a specified precision in a political opinion poll is established using the DeMoivre-Laplace Limit Theorem. This is then related via the Central Limit Theorem to the lowest number of people required in a survey measurement of a wealth-dependent parameter at the same precision. A degenerate lower limit is found for the ratio of the VPF minimum sample size to that of a political opinion poll at the same level of precision. This limiting “numbers ratio” is found to be the same under both models for the VPF survey.

Section 5 contains the discussion and Section 6 the conclusions.

Appendix A derives the mean and variance when data for cumulative probability are given in tabular form. Appendix B introduces the DeMoivre-Laplace Limit Theorem to explain the margin of error used in political opinion polls. Appendix C sets out how, for an opinion poll, the confidence interval results from a minimisation procedure applied to the interval containing the population average. Appendix D introduces the Central Limit Theorem as a generalisation of the DeMoivre-Laplace theorem and explains the derivation of a confidence interval for the population average of a continuous parameter.

The final Appendix E, details how a random multiplier may be incorporated into the Utility Model to produce the Extended Utility Model, thereby allowing for the likely variation in stated personal VPF amongst individuals with similar personal wealth.

2. The challenge of measuring a wealth-dependent parameter by survey

Figures from the UK's Office for National Statistics (ONS) show people's wealth varies over a very wide range [43]. While a small fraction, between 1% and 2%, of UK households have negative assets, the richest person in Britain is said to be worth £21bn [44]. Other developed countries are likely to exhibit similarly large variations in the individual assets of their citizens.

Approximate values for individual wealth percentiles may be found by dividing the household wealth percentiles provided by the Office of National Statistics (Fig. 3 of [43]) by the average number of people per household, namely 2.4 [45]. See Table 1 below. A small fraction (<2%) of the UK population was living with the burden of net debt in 2014–2016, but the average personal wealth was £198,112 (a figure found by dividing the ONS figure for total wealth, £12.8 tn, by the 2014 UK population of 64.6M). Meanwhile, slightly more than 10% of UK citizens had assets of £0.5M or above, and more than 1 in 50 were millionaires. The cumulative probability for individual wealth in the UK is illustrated in Fig. 1, while Fig. 2 gives the probability density for wealth in the UK in 2014–2016. The irregularity of the probability density is striking, showing a very considerable variation within small ranges as well as over the full span.

The starting wealth in Table 1 (a debt of £4933) is found by extrapolation from the ONS data. The final point is estimated after assuming that the probability density is uniform between percentiles, as in Appendix A. Using Eq. (A.5), £1,796,407 emerges as the necessary balancing wealth needed to ensure that the average wealth from Table 1 matches the figure of £198,112 calculated in the last paragraph. This provides an approximate match to the

rising trend observed in the preceding percentiles (Fig. 3). Of course, given the £21 bn of wealth ascribed to one UK citizen and the number of other known billionaires in the population, the limiting wealth is clearly a notional figure only. Its relatively low value, less than £2M, suggests that the variance on wealth will be understated as a result.

Even so, the great variation in wealth across the nation remains evident, and this poses a significant challenge to those hoping to measure, by opinion survey, the average of a wealth-dependent parameter. To put the challenge in context, the ONS felt it necessary in their wealth survey to set the size of their sample at 18,000. This sample size is very much higher than those used in most political opinion polls, for example, which tend to be in the low thousands.

The paper will examine the VPF as an exemplar of a continuous parameter that is strongly conditioned by the wealth of the individual making the response.

3. Models for a survey measurement of the VPF

This Section will derive two possible mathematical models of the survey measurement of the VPF. The first of these will provide a justification for the assumption that an individual's personal VPF is heavily dependent on his or her wealth. The second provides a general, and hence transferable, representation of a parameter that is strongly wealth dependent.

3.1. Utility Model of VPF survey

The first approach accords with that presented by Carthy et al. [33] as analysed by Thomas and Vaughan (Appendix A of [34]). Under this model the person will be prepared to sacrifice some of his/her wealth to avoid injury as long as his/her expected utility stays constant.

Let w_i be the starting wealth of the i th individual and $U_i(w_i)$ be his/her utility of wealth when he/she is in good health. Now consider an injury, k , that is assumed to reduce the individual's utility of wealth to $I_{ki}(w_i)$. Let the new utility be a fraction, γ_{ki} , of the person's utility of wealth when in good health:

$$I_{ki}(w_i) = \gamma_{ki} U_i(w_i) \quad 0 \leq \gamma_{ki} < 1 \quad (1)$$

(This equation, attributed to Viscusi and Evans [46], differs from the form, $I_{ki}(w_i) = U_i(w_i) - a_{ki}$, $a_{ki} \geq 0$, proposed by Carthy et al. [33], but those later authors clearly regarded it as a legitimate alternative (see Note 5 of [33]).)

Now suppose that, over some period of time, an individual faces two mutually exclusive possibilities: (i) he/she may incur injury k , with probability q_k , or (ii) he/she may continue in full health. (The simplifying assumption is made that the probability of death from causes other than the specified injury is negligible.) The individual's utility of wealth will now depend on whether or not the injury occurs and hence will be a random variable, $Z_i(w_i)$, with an expected value, z_i , given by:

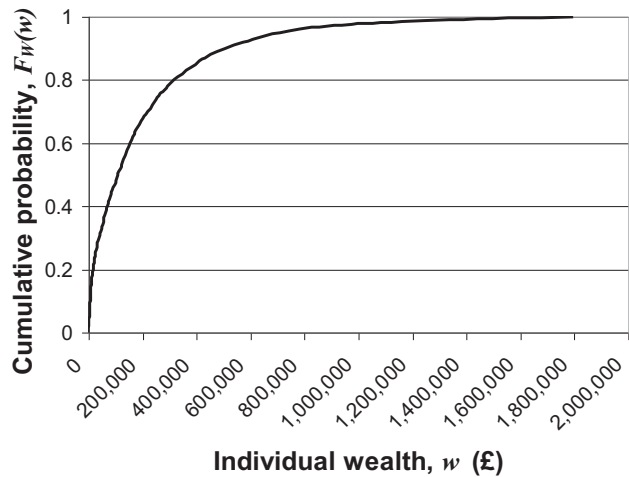
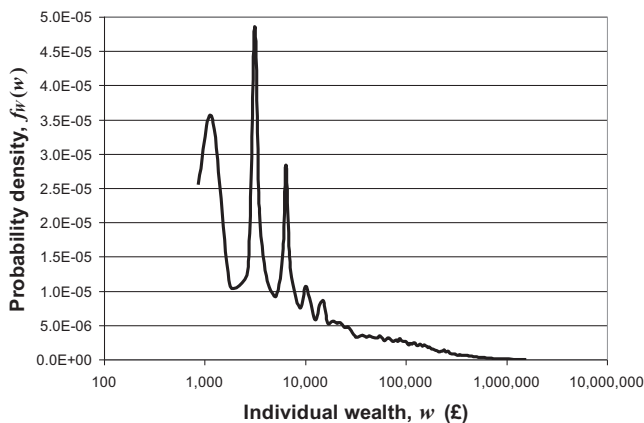
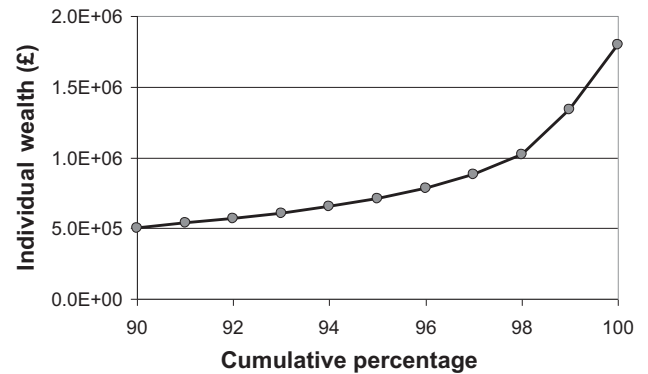
$$z_i = E[Z_i(w_i)] = (1 - q_k) U_i(w_i) + q_k I_{ki}(w_i) \quad (2)$$

It may be possible for the individual to reduce the probability, q_k , of injury, k , over some period by expending money on protection, leading to a fall in his/her wealth. Alternatively, he/she might accept a higher injury probability but receive compensation, with the result that his/her wealth, w_i , rises. Thus changes to q_k will be accompanied by related changes in w_i , and the limiting condition occurs when the expected utility stays the same, viz. $z_i = \text{constant}$, which implies that $\partial z_i / \partial q_k = 0$.

Carrying out the necessary partial differentiation of Eq. (2) with respect to q_k and setting the result to zero gives the rate of change,

Table 1Cumulative probability distribution for individual wealth, $F_W(w)$, for the UK, 2014–2016.

w (£)	$F_W(w)$	w (£)	$F_W(w)$	w (£)	$F_W(w)$	w (£)	$F_W(w)$
–4933	0.00	25,402	0.25	108,096	0.5	260,520	0.75
–2137	0.01	27,482	0.26	112,412	0.51	269,062	0.76
658	0.02	29,889	0.27	116,835	0.52	279,520	0.77
1050	0.03	32,974	0.28	120,940	0.53	291,118	0.78
1333	0.04	35,949	0.29	125,008	0.54	303,333	0.79
2267	0.05	38,738	0.3	129,792	0.55	315,170	0.8
3042	0.06	41,858	0.31	134,703	0.56	330,509	0.81
3248	0.07	44,794	0.32	139,167	0.57	345,549	0.82
3714	0.08	47,837	0.33	143,842	0.58	361,657	0.83
4586	0.09	50,981	0.34	149,035	0.59	377,527	0.84
5668	0.1	54,125	0.35	153,799	0.6	394,168	0.85
6285	0.11	57,013	0.36	158,626	0.61	411,060	0.86
6638	0.12	60,240	0.37	163,713	0.62	431,554	0.87
7296	0.13	63,875	0.38	169,613	0.63	452,598	0.88
8291	0.14	67,032	0.39	175,092	0.64	476,581	0.89
9625	0.15	70,330	0.4	181,418	0.65	504,161	0.9
10,562	0.16	73,996	0.41	187,830	0.66	536,670	0.91
11,679	0.17	77,729	0.42	195,053	0.67	571,766	0.92
13,410	0.18	81,223	0.43	202,605	0.68	607,089	0.93
14,637	0.19	85,021	0.44	210,753	0.69	654,886	0.94
15,813	0.2	88,275	0.45	219,920	0.7	709,013	0.95
17,708	0.21	91,854	0.46	227,864	0.71	780,563	0.96
19,513	0.22	95,579	0.47	236,014	0.72	881,842	0.97
21,394	0.23	99,358	0.48	242,878	0.73	1,021,726	0.98
23,259	0.24	103,307	0.49	251,503	0.74	1,336,894	0.99
						1,796,407	1.00

**Fig. 1.** UK individual wealth, 2014–2016.**Fig. 2.** Probability density for individual wealth in the UK, 2014–2016. (Note the logarithmic scale for the horizontal axis, required to allow discrimination at lower levels of wealth.)**Fig. 3.** Calculated individual wealth percentiles.

$\partial w_i / \partial q_k$, of wealth, w_i , with the probability, q_k , of injury k , for a constant expected utility

$$\frac{\partial z_i}{\partial q_k} = \left((1 - q_k) \frac{dU_i}{dw_i} + q_k \frac{dI_{ki}}{dw_i} \right) \frac{\partial w_i}{\partial q_k} + I_{ki} - U_i = 0 \quad (3)$$

Now the probability, p_k , of not receiving injury k is $p_k = 1 - q_k$, so that $dp_k/dq_k = -1$ and $\partial w_i / \partial p_k = -\partial w_i / \partial q_k$. It is reasonable to assume that a person would be prepared to trade some of his/her wealth, w_i , for a higher probability of not receiving injury k . An indifference curve should then exist, where the person's utility stays constant. The expression, $-\partial w_i / \partial p_k|_{z_i=\text{const}}$, quantifies the trade-off at a general point, (p_k, w_i) , on the indifference curve: wealth is decreased by a positive small amount, δw_i , from w_i to $w_i - \delta w_i$ in order that non-injury probability, p_k , should increase to $p_k + \delta p_k$, where δp_k is small and positive. The negative of the partial differential of wealth with respect to non-injury probability may be described as the marginal rate of substitution, m_{ki} , of non-injury probability, p_k , in place of wealth, w_i : $m_{ki} = -\partial w_i / \partial p_k|_{z_i=\text{const}}$. Since $\partial w_i / \partial p_k = -\partial w_i / \partial q_k$, it follows that

$$m_{ki} = \frac{\partial w_i}{\partial q_k} \Big|_{z_i=\text{const}} = \frac{U_i(w_i) - I_{ki}(w_i)}{(1 - q_k) \frac{dU_i}{dw_i} + q_k \frac{dI_{ki}}{dw_i}} = \frac{(1 - \gamma_{ki})U_i(w_i)}{(1 - q_k(1 - \gamma_{ki})) \frac{dU_i}{dw_i}} \quad (4)$$

It is argued in [14,15] that only utility functions from the Power family can be true descriptions of human experience. Hence the utility of wealth may be written:

$$U_i(w_i) = \begin{cases} \alpha_i \frac{w_i^{1-\varepsilon} - 1}{1-\varepsilon} & \text{for } \varepsilon \neq 1 \\ \alpha_i \ln w_i & \text{for } \varepsilon = 1 \end{cases} \quad (5)$$

where ε is risk-aversion and a constant. α_i is also a constant, conforming with the ideas of von Neumann and Morgenstern [13] and the clarification brought out in [14,15] that the Power utility formulation of Eq. (5) gives the utility relative to the utility of one unit of wealth, thus producing a ratio scale for utility of wealth.

Differentiating equation pair (5) with respect to wealth gives:

$$\frac{dU_i}{dw_i} = \frac{\alpha_i}{w_i^\varepsilon} \quad \text{for all } \varepsilon \quad (6)$$

Thus, substituting from Eqs. (5) and (6) into Eq. (4) gives:

$$m_{ki} = \begin{cases} \frac{1-\gamma_{ki}}{1-q_k(1-\gamma_{ki})} \frac{w_i - w_i^\varepsilon}{1-\varepsilon} & \text{for } \varepsilon \neq 1 \\ \frac{1-\gamma_{ki}}{1-q_k(1-\gamma_{ki})} w_i \ln w_i & \text{for } \varepsilon = 1 \end{cases} \quad (7)$$

Now let k denote a fatal injury, in which case $m_{ki} = m_{Di} = v_i$, the marginal rate of substitution of non-death probability in place of wealth, which is individual i 's personal VPF:

$$v_i = \begin{cases} \frac{1-\gamma_{Di}}{1-q_D(1-\gamma_{Di})} \frac{w_i - w_i^\varepsilon}{1-\varepsilon} & \text{for } \varepsilon \neq 1 \\ \frac{1-\gamma_{Di}}{1-q_D(1-\gamma_{Di})} w_i \ln w_i & \text{for } \varepsilon = 1 \end{cases} \quad (8)$$

Eq. (8) is dependent both on both the individual's fractional loss, $1 - \gamma_{Di}$, of utility through dying prematurely and on the probability of being killed, q_D . However, we may judge that that nearly all utility is lost on death for the average person, so that $\gamma_{Di} \rightarrow 0$. Presumably γ_{Di} will attain exactly that figure for an individual without close relatives and uninterested in charitable giving. Moreover in many cases, for example when an improvement is being considered for an industrial protection system, the probability of death is already low: $q_D \ll 1$. Hence equation set (8) may be approximated by:

$$v_i \approx \begin{cases} \frac{w_i - w_i^\varepsilon}{1-\varepsilon} & \text{for } \varepsilon \neq 1 \\ w_i \ln w_i & \text{for } \varepsilon = 1 \end{cases} \quad (9)$$

A significant failing of the model just presented is the lack of any dependence on age even though what is at stake for a typical young person is patently very different from that for an elderly individual – the VPF values the life of a 20 year-old with 61 years' expected life the same as that of a 90 year-old with about 4 years of life expectancy. This reflects a major philosophical weakness in the VPF concept, as been pointed out previously by Sunstein [47] as well as Thomas and Vaughan [32].

By the theory just presented, and from Eq. (9) in particular, the individual's VPF is determined entirely by his or her wealth as soon as the risk-aversion, ε , is specified, and randomness is introduced only by the process of selecting a person from the population. The absence of variation in stated VPFs amongst people with the same wealth is clearly a simplification that will reduce the variance predicted by the model. (This issue is explored in Appendix E.)

A table of percentiles for the VPF may be built up by applying equation set (9) to the percentiles of wealth given in Table 1. The methods of Appendix A may then be applied to find the population average VPF and the variance of personal VPF.

The first two percentiles of wealth are negative, which poses a problem in computing the corresponding figures for VPF from Eq. (9). Two potential ways forward may be considered. In the first, $|w_i|$ and $|w_i|^\varepsilon$ are substituted in place of w_i and w_i^ε respectively in Eq. (9) and the utility of negative wealth is calculated as a negative

multiple, $-a : a \approx 2$, of the result. This takes up the notion, introduced by Kahneman and Tversky [48] and discussed further in [49], that the disutility of a debit is numerically greater than the utility of the same sum in credit. Hence

$$v_i = \begin{cases} |w_i|^\varepsilon \times -a \frac{|w_i|^{1-\varepsilon} - 1}{1-\varepsilon} & \text{for } \varepsilon \neq 1; i = 1, 2 \\ |w_i| \times -a \ln |w_i| & \text{for } \varepsilon = 1; i = 1, 2 \end{cases} \quad (10)$$

which produces negative values for v_1 and v_2 . The other approach is simply to set $v_1 = v_2 = 0$. In fact it matters little which of these methods is used, as the changes in calculated values for μ_V and σ_V^2 are reflected only in the 4th significant figure. Hence it is reasonable to choose the simpler second option. Such a course avoids the awkward implication that those in debt will regard death as a financially attractive option, although it can still be inferred that they regard death as financially neutral.

The validation of the J-value method [20] means that it is possible to find a good estimate of the value to be placed on the average life to come for UK citizens. Carrying out this exercise for actuarial and GDP data for 2015 gives a figure of £8,587,700. Taking this to be the actual value of the VPF, an implied value for ε may be found, namely that which causes the expected value, $E(V)$, to coincide with this figure. $E(V)$ may be found by calculating the personal VPF, v_i , for each wealth, w_i , listed in Table 1 and applying Eq. (A.5) from Appendix A. The risk-aversion is then varied to give a match, which occurs at a risk-aversion of $\varepsilon = 1.1605$. Thus the deduced value for risk-aversion lies in the second quartile of the range, $1 \leq \varepsilon \leq 1.5$, currently advocated in the UK Treasury's Green Book [50]. It conforms also to the Treasury's previous advice in its Green Book, namely that ε was likely to be just below or just above unity [51].

Applying Eq. (A.9) with Eqs. (A.5) and (A.6) of Appendix A, and putting $X = V$, $x_i = v_i$ and $n = 100$, allows the variance, σ_V^2 , to be found. Eq. (11) summarises the key figures coming from the Utility Model:

$$\begin{aligned} \mu_V &= £8,587,700 \\ \sigma_V^2 &= 1.8299 \times 10^{14} \\ \sigma_V &= £13,527,463 \end{aligned} \quad (11)$$

3.2. Multiplier Model of VPF survey

An alternate model assumes that the individual's stated VPF is some multiple, R , of his/her wealth: the "Multiplier Model". Rather than assigning the multiplier the same value for all people, it is assumed that R is a random variable, independent of individual wealth. This provides for likely variations in the responses of people with similar wealth, allowing for the presence of other, unspecified influences such as differences in psychological makeup.

The individual's VPF, V , is then regarded as a continuous random variable obeying the relation:

$$V = RW \quad (12)$$

where W is wealth, with W and R taken to be continuous random variables, independent the one of the other. Let the multiplier, R , take the form:

$$R = r_0 + Q \quad (13)$$

where $r_0 \geq 0$ is deterministic and Q is a random variable distributed uniformly over the interval, $(0, s)$. Thus the probability density function, $f_Q(q)$, for Q is

$$f_Q(q) = \frac{1}{s} \quad 0 \leq q \leq s \quad (14)$$

The expected value of the multiplier is therefore:

$$E(R) = E(r_0 + Q) = r_0 + E(Q) = r_0 + \frac{s}{2} \quad (15)$$

based on the standard result that the expected value for a uniform distribution is its central value. The maximum value, $r_{0\max}$, of the deterministic component, r_0 , is that which causes the random band to degenerate to a point of zero length, when $E(R) = r_{0\max}$.

Given that R and W are modelled as independent, the expected value of their product is simply the product of their expected values:

$$E(V) = E(R)E(W) \quad (16)$$

The average value of individual wealth is calculated using Eq. (A.5) of Appendix A as:

$$E(W) = £198,112 \quad (17)$$

As was the case with the Utility Model, $E(V)$ is set to the monetary value assigned to population-average life expectancy by J-value analysis, namely £8,587,700. Hence from Eq. (16), the expected value of the multiplier, R , is:

$$E(R) = \frac{E(V)}{E(W)} = \frac{£8,587,700}{£198,112} = 43.35 \quad (18)$$

Using Eq. (15), the extent of the random band, s , is then

$$s = 86.7 - 2r_0 \quad (19)$$

Since r_0 is deterministic, the variance of R is found by applying the variance operator to Eq. (13), revealing it to be simply the variance of Q :

$$\text{var}(R) = \text{var}(r_0 + Q) = \text{var}(Q) = \frac{s^2}{12} \quad (20)$$

where the last step uses a standard result for a uniform distribution.

Meanwhile the variance of individual wealth may be found from the tabulated data of Table 1 using Eq. (A.9) of Appendix A:

$$\text{var}(W) = £6.8380 \times 10^{10} \quad (21)$$

The variance of V is the product of the two independent random variables, R and W , which may be written [52]:

$$\text{var}(V) = \text{var}(RW) = \text{var}(R)\text{var}(W) + (E(W))^2\text{var}(R) + (E(R))^2\text{var}(W) \quad (22)$$

with the square root giving the associated standard deviation: $\sigma_V = \sqrt{\text{var}(V)}$.

Varying the deterministic component, r_0 , of the random multiplier, R , will cause the length, s , of the random band to vary. At the upper limit for r_0 , $r_0 = r_{0\max}$, the multiplier will be wholly deterministic, while at the lower end of the range, $r_0 = 0$, the multiplier, R , will be uniformly distributed over the full range, $0 \rightarrow 2r_{0\max}$. Fig. 4 shows the effect on the probability distribution, $f_Q(q)$, for the random component, Q , of the multiplier as deterministic component, r_0 , will be reflected in alterations to the standard deviation, σ_V , for the VPF. Fig. 5 shows the standard deviation, σ_V , plotted against the value selected for the deterministic component, r_0 , over the range of possible values: $0 \leq r_0 \leq r_{0\max}$, where $r_{0\max} = E(R) = 43.35$.

The standard deviation for VPF takes its highest value, £14M when the multiplier, R , is uniformly random between 0 and $2E(R)$, viz. between 0 and 86.7. Most of the variability comes from the variation of individual wealth in the population, as can be seen from the last point on the graph, where the multiplier is fixed at $r_0 = r_{0\max} = E(R) = 43.35$ and has zero variance. Putting $\text{var}(R) = 0$ in Eq. (22) then gives

$$\text{var}(V) = (E(R))^2\text{var}(W) = r_{0\max}^2\text{var}(W) \quad (23)$$

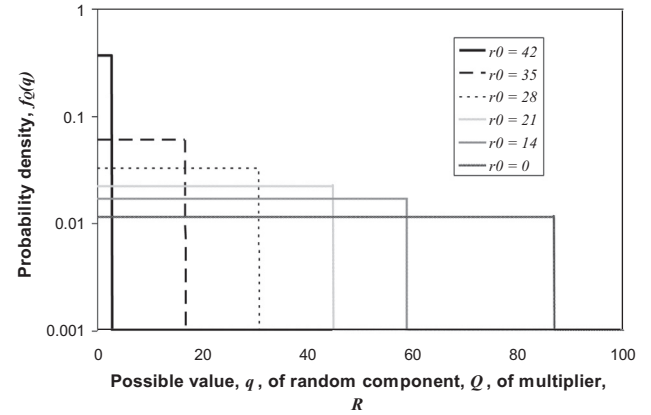


Fig. 4. Probability density for the random component, Q , of the random multiplier, R , with the size of deterministic element, r_0 , as parameter.

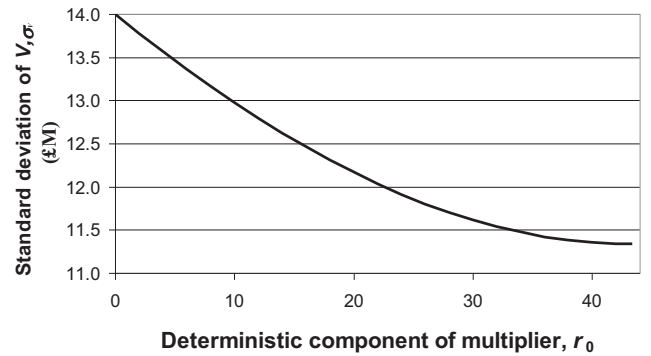


Fig. 5. Dependence of the standard deviation of the VPF on the deterministic component of the multiplier, r_0 . (Multiplier model).

At this point $\sigma_V = r_{0\max}\sigma_W = £11.34M$. This is only 19% down on the £14M figure for standard deviation of the VPF that is associated with a maximally random model for R .

4. Comparison with political opinion polls

4.1. The numbers of respondents in a political opinion poll and in a survey measurement of a general continuous parameter, A

For compactness, the term, “ β -level precision” is introduced in this paper as a measure of the precision afforded by the average value derived from a survey. It is taken to denote the length of the $\beta\%$ confidence interval for the population mean divided by the best available estimate of that mean.

Appendix B examines the mathematical representation of political opinion polls and sets out the relationships amongst the confidence interval, the margin of error and the β -level precision. Appendix D explains the relation between the β -level precision and the confidence interval for a general, continuous parameter measured by survey.

The β -level precision, $\lambda_A^{(\beta)}$, for the measurement of a general continuous variable, A , is found from Eq. (D.9) of Appendix D. Particularising σ , μ and n by assigning the subscript, A , gives:

$$\lambda_A^{(\beta)} = 2b^{(\beta)} \frac{\sigma_A}{\mu_A \sqrt{n_A}} \quad (24)$$

where $b^{(\beta)}$ is the solution to implicit Eq. (B.11) in Appendix B. Thus, for example, at confidence level of 90%, $b^{(0.90)} = 1.645$, while at a confidence level of 95%, $b^{(0.95)} = 1.96$.

$\lambda_A^{(\beta)}$ will be an increasing function of β , so that, for example, $\lambda_A^{(0.95)} > \lambda_A^{(0.9)} > \lambda_A^{(0.8)}$. Once β is specified, $\lambda_A^{(\beta)}$ provides a quantitative measure of the precision of the survey measurement of A. At any given β , the smaller the value of $\lambda_A^{(\beta)}$, the better the precision.

Take now the case of an opinion poll aimed at establishing the fraction of the population holding a certain view in an upcoming, tight vote between alternatives. In such a case, β -level precision takes the simple form (see Appendix B, Eq. (B.20)):

$$\lambda_{poll}^{(\beta)} = \frac{2b^{(\beta)}}{\sqrt{n_{poll}}} \quad (25)$$

where n_{poll} is the number of respondents to the opinion poll.

Let us now compare two different survey measurements, one attempting to measure a continuous parameter, A, and the other a political opinion poll. If they exhibit the same β -level precision, that is to say if $\lambda_A^{(\beta)} = \lambda_{poll}^{(\beta)}$, then, from Eqs. (24) and (25):

$$\frac{n_A}{n_{poll}} = \frac{\sigma_A^2}{\mu_A^2} = \frac{\text{var}(A)}{(E(A))^2} \quad (26)$$

Two features of Eq. (26) are noteworthy:

- (i) if the survey measurement of a parameter, A, and the political opinion poll are to have the same β -level precision, the ratio of the number of respondents in the survey for A to the number in the political opinion survey will depend solely on the distribution of A in the population
- (ii) the result applies to all confidence levels, β .

4.2. The margin of error and the “12% rule”

Most people are familiar with opinion pollsters quoting a “margin of error” to indicate the likely precision of their political predictions. The margin of error may be defined as half the length of the $\beta=95\%$ confidence interval when a binary choice is given and the probability of either option being supported in the actual vote is about 50% (see Appendices B and C). Such conditions were met, for example, in the UK’s EU Referendum of 2016 [54], where 51.9% of the voting public chose the option of leaving the European Union while 48.1% wanted to remain.

Political opinion pollsters have reached a *de facto* consensus in the UK that the margin of error should be no greater than 3%. This figure implies that the length, w , of the 95% confidence interval should be no more than 6% in absolute terms. Referred to the fraction of the people voting for a particular option of about 0.5 when the vote is close, this implies a β -level precision, $\lambda_{poll}^{(\beta)}$, of 12% or better for a political opinion poll with $\beta = 95\%$: $\lambda_{poll}^{(0.95)} \leq 0.12$. Such a β -level precision may be seen as corresponding to a β -level tolerance of $\pm 6\%$ or better at $\beta = 95\%$.

It seems reasonable to require that a survey estimate of any important population parameter, A, should match the precision of the political opinion polls frequently carried in newspapers and the broadcast media. This implies that the β -level precision at $\beta = 95\%$ should be 12% or better for the survey measurement of A.

$$\lambda_A^{(0.95)} = \lambda_{poll}^{(0.95)} \leq 0.12 \quad (27)$$

This criterion may be called the “12% rule”.

When a political opinion poll obeys the 12% rule, then applying Eq. (B.20) of Appendix B to condition (27) gives:

$$\frac{3.92}{\sqrt{n_{poll}}} \leq 0.12 \quad (28)$$

so that

$$n_{poll} \geq \left(\frac{3.92}{0.12} \right)^2 \quad (29)$$

Thus $n_{poll} \geq 1068$ after allowing for the integer requirement.

The sample size for the VPF for the same β -level precision is given, using Eq. (26), as

$$n_V = \frac{\text{var}(V)}{(E(V))^2} n_{poll} \quad (30)$$

For the Utility Model for a VPF survey, the numerical values of $E(V)$ and $\text{var}(V)$ are given in Eq. (11), so that, combining Eq. (30) with inequality (29), the sample size when the 12% rule is obeyed is

$$n_V \geq 2648 \quad (31)$$

A range of values of $E(V)$ and $\text{var}(V)$ are possible for the Multiplier Model for VPF survey, corresponding to different values of the fixed component, r_0 , of the multiplier, R . Allowing for r_0 to span its full range from 0 to $r_{0\max} = 43.35$ gives the following interval for the minimum sample number, $n_{V\min}$, that will obey the 12% rule and thus provide a β -level precision of $\lambda_V(0.95) \leq 0.12$:

$$1860 \leq n_{V\min} \leq 2835 \quad (32)$$

Fig. 6 graphs the minimum sample size, $n_{V\min}$, when the 12% rule is obeyed for the Multiplier Model for VPF. The ratio of the fixed component, r_0 , of multiplier, R , ranges from 0.0, where R is fully stochastic, to $r_{0\max}$, where the multiplier becomes entirely deterministic. Also marked up is the minimum number of respondents needed by the Utility Model of VPF survey when the 12% rule is obeyed.

Comparing inequalities (31) and (32), it is clear that the two models for VPF survey produce similar estimates for minimum sample size, with the figure for the Utility Model falling within the range set by the Multiplier Models.

It may be concluded that the sample size for a VPF survey needs to be set much higher than for a political opinion poll if both surveys are to satisfy the 12% rule.

4.3. Degenerate lower limits of the “numbers ratio” of the VPF minimum sample size to the minimum number of opinion poll respondents at the same β -level precision

4.3.1. The numbers ratio n_V/n_{poll} for the Utility Model for VPF survey

For the case where $\varepsilon \neq 1$, the VPF is given by the Utility Model as:

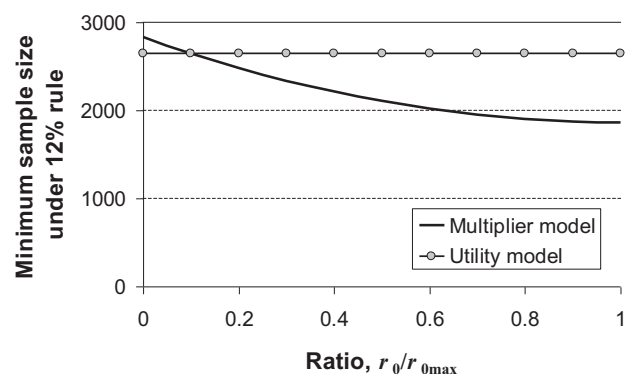


Fig. 6. Minimum sample size, $n_{V\min}$, under the 12% rule for the multiplier and Utility Models for VPF survey.

$$V = \frac{1}{\varepsilon - 1} (W^\varepsilon - W) \quad \varepsilon \neq 1 \quad (33)$$

Applying the expectation operator, $E(\cdot)$, gives

$$E(V) = \frac{1}{\varepsilon - 1} (E(W^\varepsilon) - E(W)) \quad \varepsilon \neq 1 \quad (34)$$

while the application of the variance operator, $\text{var}(\cdot)$, to Eq. (33) produces:

$$\begin{aligned} \text{var}(V) &= \left(\frac{1}{\varepsilon - 1}\right)^2 \text{var}(W^\varepsilon - W) \\ &= \left(\frac{1}{\varepsilon - 1}\right)^2 (\text{var}(W^\varepsilon) + \text{var}(W) - 2\text{cov}(W^\varepsilon, W)) \end{aligned} \quad \varepsilon \neq 1 \quad (35)$$

But

$$\text{cov}(W^\varepsilon, W) = \rho \sqrt{\text{var}(W^\varepsilon)} \sqrt{\text{var}(W)} \quad (36)$$

where ρ is the correlation coefficient. When risk-aversion is $\varepsilon = 1.1605$, the correlation coefficient is 0.9955 for the range of wealths listed in Table 1. See Fig. 7.

Combining Eq., (34)–(36) gives:

$$\begin{aligned} \frac{\text{var}(V)}{(E(V))^2} &= \frac{\text{var}(W^\varepsilon) + \text{var}(W) - 2\rho \sqrt{\text{var}(W^\varepsilon)} \sqrt{\text{var}(W)}}{(E(W^\varepsilon) - E(W))^2} \\ &\approx \frac{\text{var}(W^\varepsilon) - 2\sqrt{\text{var}(W^\varepsilon)} \sqrt{\text{var}(W)} + \text{var}(W)}{(E(W^\varepsilon) - E(W))^2} \end{aligned} \quad \varepsilon \neq 1 \quad (37)$$

where $\rho \approx 1$ has been used in the second step. Factorising the numerator as a square and using Eq. (26), the ratio of the sample size for a VPF survey to that for a political opinion poll giving the same β -level precision is:

$$\frac{n_V}{n_{\text{poll}}} = \frac{\text{var}(V)}{(E(V))^2} \approx \left(\frac{\sqrt{\text{var}(W^\varepsilon)} - \sqrt{\text{var}(W)}}{E(W^\varepsilon) - E(W)} \right)^2 \quad (38)$$

Since the distribution of W^ε is conditioned by the distribution of wealth, W , Eq. (38) shows that, under the Utility Model of VPF survey, the ratio of numbers of VPF respondents and political poll respondents giving the same β -level precision for any β , depends only on the distribution of wealth.

Fig. 8 plots the numbers ratio, $n_V/n_{\text{poll}} = \text{var}(V)/(E(V))^2$, against risk-aversion, ε , when the political opinion poll and the survey measurement of VPF have a common β -level precision. (The point at $\varepsilon = 1$ may be found numerically as $\lim_{\varepsilon \rightarrow 1} \text{var}(V)/(E(V))^2$, with $\varepsilon = 1 \pm \delta$; $\delta = 1.0 \times 10^{-8}$.)

The numbers ratio, n_V/n_{poll} , may be seen to reach a minimum over all positive risk-aversions: $\varepsilon \geq 0$ at the point where $\varepsilon = 0$.

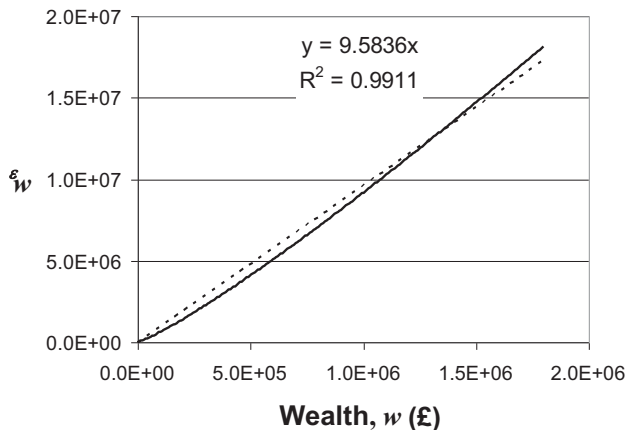


Fig. 7. Correlation between w^ε and w .

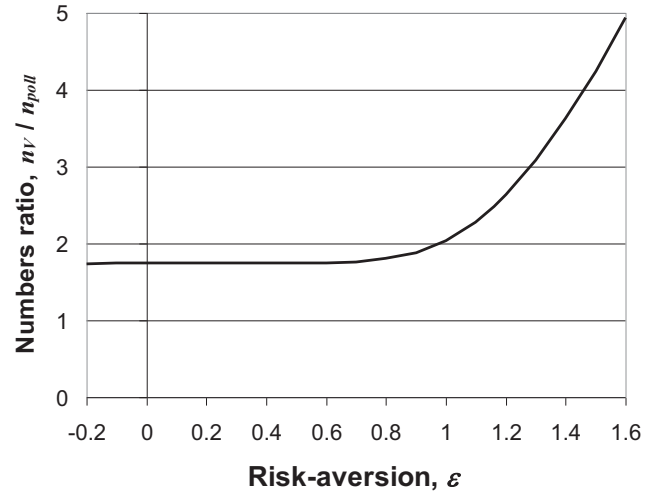


Fig. 8. Numbers ratio, n_V/n_{poll} , for Utility Model of VPF survey and opinion poll, at the same β -level precision.

The decrease in the numbers ratio continues as the risk-aversion falls below 0.0 and moves into the risk confident region, but the difference from the value at $\varepsilon = 0$ affects only the 6th significant figure as ε decreases over the interval, $0 \rightarrow -0.5$.

There is a limiting value of the ratio, $\text{var}(V)/(E(V))^2$, for the degenerate case when the risk-aversion is zero. Inserting $\varepsilon = 0$ into either Eq. (38) or the first line of Eq. (37) to give:

$$\frac{\text{var}(V)}{(E(V))^2} = \frac{\text{var}(W)}{(E(W) - 1)^2} \approx \frac{\text{var}(W)}{(E(W))^2} \quad (39)$$

where the last step follows from the fact that $E(W)$, at £198,112, is 5 orders of magnitude greater than £1.

Each person's individual VPF will decrease as ε drops, and by Eq. (9), when $\varepsilon = 0$, then $v_i = w_i - 1 \approx w_i$. Hence the average value of the VPF will fall to approximately the average value of wealth, £198,112 (strictly to £198,111).

On combining Eqs. (39) and (30), it is clear that, no matter what value of risk-aversion is employed in the utility function (and hence whatever the VPF produced), the number of respondents needed to measure the VPF by survey must be at least 74% bigger than the number consulted in a political opinion poll giving the same β -level precision.

It will be shown in Section 4.3.2 that the Multiplier Model produces the same degenerate limiting value for the numbers ratio, n_V/n_{poll} , when the β -level precision is the same for the VPF survey as for a political opinion poll.

4.3.2. The numbers ratio n_V/n_{poll} for the Multiplier Model for VPF survey

Dividing Eq. (22) by $(E(V))^2$ and using Eq. (16) produces

$$\begin{aligned} \frac{\text{var}(V)}{(E(V))^2} &= \frac{\text{var}(R)\text{var}(W) + (E(W))^2 \text{var}(R) + (E(R))^2 \text{var}(W)}{(E(R))^2 (E(W))^2} \\ &= \frac{\text{var}(R)}{(E(R))^2} \left(1 + \frac{\text{var}(R)}{(E(R))^2} \right) + \frac{\text{var}(R)}{(E(R))^2} \end{aligned} \quad (40)$$

Hence

$$\frac{\text{var}(V)}{(E(V))^2} \rightarrow \frac{\text{var}(W)}{(E(W))^2} \quad \text{as } \text{var}(R) \rightarrow 0 \quad (41)$$

In the degenerate case, as $r_0 \rightarrow r_{0\text{max}}$, so the variance, $\text{var}(R)$, disappears and the multiplier becomes entirely deterministic. Eq. (41) is the outcome, a result for the Multiplier Model that mirrors Eq. (39) for the Utility Model when the risk-aversion is set to zero.

Eq. (41) will be valid for all $E(R) : E(R) \neq 0$ and thus will apply to any positive expected value, $E(R)$, set for the multiplier. This conclusion holds irrespective of the initial probability distribution for the multiplier, R . The resulting condition gives the numbers ratio as:

$$\frac{n_V}{n_{poll}} \geq \frac{\text{var}(W)}{(E(W))^2} = 1.74 \quad (42)$$

if the β -level precision of the VPF survey is to match that of the political opinion poll.

The fact that there is no dependence on $E(R)$ means that Eq. (42) will hold whether the VPF is assumed to be £8.5M or £1M or £2M or £20M or any other value.

The generality of these arguments means that they may be applied to the survey measurement in the UK of any parameter, A , that is strongly wealth-dependent. Hence the numbers ratio will obey the condition:

$$\frac{n_A}{n_{poll}} \geq 1.74 \quad (43)$$

when the β -level precision for the survey for parameter, A , is the same as that of the opinion poll. Condition (43) devolves into an equality when the measurand, A , is proportional to wealth; such an eventuality includes the case where what is to be measured is wealth itself. (This last statement follows directly from Eq. (26), of course.)

4.4. β -Level precisions at $\beta = 95\%$ for different sample sizes

Fig. 9 gives another perspective by plotting, against sample size, the β -level precisions for the Utility and the Multiplier Models for VPF survey, as well as the β -level precision for a political opinion poll, all at $\beta = 95\%$. The curve for the Multiplier Model of VPF survey incorporates the assumption that r_0 takes its central value, $\frac{1}{2}r_{0\max}$.

It is clear, when the number of respondents is the same, that the precision of the opinion poll prediction will be significantly better than that of either model of VPF survey. Conversely, the Utility and Multiplier Models for VPF survey require a significantly greater sample size to achieve the same level of precision as a political opinion poll.

It is also evident that the β -level precision of the VPF survey will deteriorate increasingly rapidly for sample sizes less than about 500, irrespective of which model of the VPF survey is used.

Fig. 9 may also be as a test for a biased sample. Suppose that a survey is carried out to measure a strongly wealth-dependent

parameter but that the number of opinions canvassed is low, say 100 people. Unbiased estimates might still be made of the mean and the standard deviation based on the opinions in the 100-strong sample. If these were used to calculate a β -level precision at $\beta = 95\%$ that was much lower than the curves in Fig. 9, at say 14%, this would suggest that the opinions surveyed showed too little diversity, implying that the sample had not been chosen randomly across the board. Achieving a β -level precision at $\beta = 95\%$ that was similar to those predicted for this sample size by the Utility or Multiplier Model would, however, be no guarantee that people with the necessary spread of wealths had been canvassed.

5. Discussion

Statistics from the UK's Office for National Statistics enable the range of wealth held by individuals in the United Kingdom to be assessed and the distribution to be characterised by the percentiles listed in Table 1. Extrapolation has been needed for the starting and finishing points (0% and 100%), and the top figure for personal wealth in the table is accepted as being too low. While Table 1 encompasses an extensive range of wealth, the variance of individual wealth in the UK will tend to be somewhat understated even so.

The size of sample needed to give adequate precision may be appreciated by reference to the "margin of error" used by political opinion surveyors. All other things being equal, the greater the sample size the better the precision of the opinion poll's prediction. General agreement appears to have emerged amongst the UK's practitioners that a margin of error of 3% is the highest figure at which a reasonable compromise may be struck between the expense of the survey and the commercially valuable reputation of the polling organisation. A margin of error of 3% or better will require the political opinion pollster to question at least 1068 randomly selected people.

The concept of " β -level precision", $\lambda^{(\beta)}$, can be used to generalise the margin of error to other types of survey. The "12% rule" extends the 3% margin of error or better ($\lambda^{(0.95)} \leq 0.12$) customary for a political opinion poll to surveys of continuous parameters. Measurement exercises conducted in line with this rule will encompass the true mean within $\pm 6\%$ of the estimated average about 95% of the time, which seems a reasonable aim.

The number of respondents needed in a survey measurement of a general continuous parameter, A , may be related to the size of the sample required for a political opinion poll when both have to satisfy the same β -level precision. The "numbers ratio" of the sample sizes in the two surveys then equals the variance of A divided by the square of its expected value. This result, valid at all values of β and β -level precision, means that the sample size for survey measurement of a continuous parameter that is strongly dependent on wealth will need always to be substantially greater than that of a political opinion poll.

Applying this finding to individual wealth in the UK gives the numbers ratio as 1.74, so that the number of people needing to be consulted in the determination of average wealth needs to be 1860 in order to just satisfy the 12% rule: $\lambda_W^{(0.95)} = 0.12$. This sample size may be compared with the equivalent number, 1068, needed for a political opinion poll at the same β -level precision, viz. $\lambda_{poll}^{(0.95)} = 0.12$.

The "value of a prevented fatality" is an example of a population parameter that is conditioned by wealth. Two diverse mathematical descriptions are offered of the process of measuring the population average VPF by opinion survey. The first, the Utility Model, accords with the description presented in the VPF paper by Carthy et al. [20], as analysed in [34]. The Utility Model provides a justification for the intuition that the individual's personal VPF will depend strongly on his/her wealth.

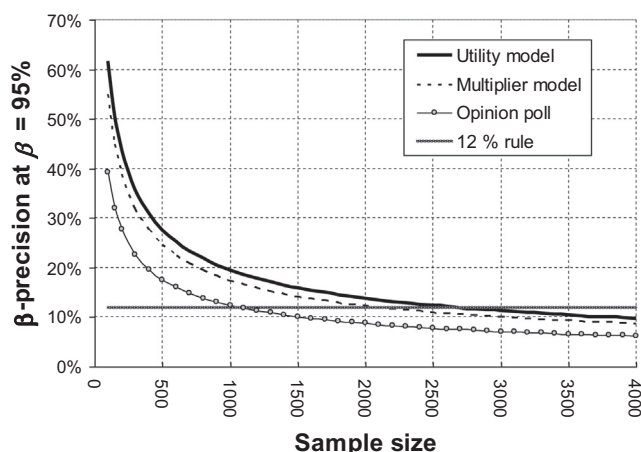


Fig. 9. β -Level precision at $\beta = 95\%$ as a function of sample size for the two models for VPF survey and for the opinion poll. The 12% rule is also marked.

Building on this confirmation, a second, conceptually simpler mathematical description is developed for the VPF survey, namely the “Multiplier Model”. Now the person’s VPF is found by multiplying his/her wealth by a random factor that is uniformly distributed between limits. Two sources of random variation are present in this model, first in the process by which a respondent is selected and second in the size of the multiplier.

The parameters of the two models may be adjusted so that each produces a population average VPF to match the average value of life to come for a UK citizen found by the J-value, viz. £8,587,700 (2015 £s). For the Utility Model of the VPF survey, this means that the risk-aversion parameter should take the value $\varepsilon = 1.1605$, which is well within ranges for risk-aversion suggested by the UK Treasury. Matching the same figure for the VPF in the Multiplier Model requires the expected value of the multiplier to be 43.35.

Similar variances for the individual’s VPF then emerge from the two independent systems of equations. There is overlap in the values of standard deviation predicted by the two models, with equality occurring at the point where the deterministic factor, r_0 , in the multiplier, R , is about 10% of its maximum value, $r_{0\max}$, in the Multiplier Model, implying that the multiplier then contains only a small deterministic component.

Under the 12% rule, the minimum sample size is 2648 for the Utility Model of the VPF survey. Meanwhile the minimum number of people needing to be consulted is estimated under the Multiplier Model to lie between 1860 and 2835, a range that encompasses the figure coming from the Utility Model.

There is a degenerate lower limit for the numbers ratio of the minimum sample size in a VPF survey to the corresponding figure in a political opinion poll at matching β -level precisions. This degenerate lower limit is the same for both the Utility Model and the Multiplier Model, and is found to depend only on the distribution of wealth in the population. It is thus clear that the need for high sample sizes is driven mainly by the variation in individual wealth in the nation.

In the case of the Utility Model of VPF survey measurement, the degenerate lower limit for minimum sample size, 1860, occurs when the risk-aversion, ε , declines to zero. The population average VPF will decrease in tandem, and will be to all intents and purposes identical to the average individual wealth, £198,000, by the time risk neutrality has been reached. At this point the survey measurement of the VPF will share the characteristics of wealth measurement by survey. This explains why the sample size has fallen to exactly that needed to find the average wealth under the 12% rule, as discussed in paragraph 5 of this Section.

The same degenerate lower limit for sample size, 1860 obtains in the Multiplier Model when the deterministic component of the multiplier, r_0 , has reached its maximum value, $r_0 = r_{0\max} = E(R)$, and the multiplier is fully deterministic. The result is independent of the value previously specified for the expectation, $E(R)$, of the multiplier, R . Thus at least 1860 people would always need to be consulted in order for the measurement to conform to the 12% rule, no matter what the value chosen for $E(R)$ and hence irrespective of the value of the VPF under the Multiplier Model.

The convergence of the two diverse models at a degenerate lower limit for minimum sample size suggests that the number of people surveyed will always need to be at least 1860, irrespective of the value of the VPF.

The next agreement between the two VPF survey models occurs at a minimum sample size of 2648 people, the figure found using the Utility Model when set up to give a VPF to match the J-value generated figure of £8.59M. This sample size is, in fact, likely to be an underestimate as it does not allow for the variation in response likely amongst people with very similar levels of wealth. Including this realistic possibility would increase the necessary

sample size. The effect has been quantified in [Appendix E](#) by extending the Utility Model through the introduction of a multiplier to represent the variation likely between the stated VPFs for people with very similar levels of wealth. Allowing for a random variation of $\pm 50\%$ on each person’s calculated personal VPF takes the minimum sample size to 2958.

Applying the 12% rule, the range of VPFs suggested by the Multiplier Model is $1860 \rightarrow 2835$, while the Extended Utility Model gives the overlapping range $2648 \rightarrow 3887$. The midpoint of the extremal values is thus $(1860 + 3887)/2 = 2874$. It is possible to conclude, in broad terms, that the minimum sample size for measuring the VPF by survey is likely to be 3000 or more randomly chosen people in order to comply with the 12% rule. This level of precision, equivalent to a tolerance of $\pm 6\%$ at the 95% confidence level, seems a reasonable objective. For comparison, this sample size is still only a sixth of the 18,000 people questioned by the Office of National Statistics in its survey of household wealth.

While the Utility Model is specific to the VPF, the Multiplier Model provides a general paradigm that can be applied to the survey measurement of any continuous parameter that is strongly dependent on wealth. In all such cases the minimum sample size must be 1860 if the 12% rule is to be obeyed, and it may well be that more people will need to be questioned.

As mentioned in the Introduction, the sample size used in the opinion survey [33] on which the UK VPF is based was only 167. [In fact Carthy et al. reduced their sample size to 149 (Table 3 of [33]) to find the set of “trimmed means” that influenced their finally recommended VPF value; meanwhile, in generating their Table 7, they reduced the number of people in their sample to 135]. It is clear from the minimum sample sizes derived in this paper and discussed above that the sample size on which the UK VPF is based was at least an order of magnitude too small, and hence wholly inadequate. Obviously any similar exercise in the future would need to be scaled up very significantly so that the views of 3000 or more randomly chosen individuals could be sampled.

[Table 2](#) shows the sample size figures for the Multiplier Model for VPF, the Utility Model and the Extended Utility Model discussed in [Appendix E](#). These are compared and contrasted with the sample size used by the ONS for measuring individual wealth [43] and the sample size used to generate the UK VPF [33]. Small scale exercises in opinion surveying to find a VPF are clearly ruled out for the future.

A further feature of the Carthy study [33] is the anomalously low level of the average wealth of its respondents, as revealed in [34]. The problem of obtaining proper representation for wealthier people in the statistics is recognised by the ONS. Its statisticians stress in the commentary on their survey of national wealth that special care is needed to include adequate number of richer respondents, saying:

“As wealth is known to be unevenly distributed, addresses more likely to contain wealthier households were sampled at a higher rate to improve the efficiency of the sample. These addresses were identified utilising data from HMRC.” – Section 8: Quality and methodology of [43].

Equal care is clearly needed in the survey of any population parameter, such as the VPF, known to be strongly dependent on wealth.

The figures for sample size derived in the paper are predicated on the assumption that the sample is chosen randomly across the whole population. As noted by the ONS, it will generally not be easy in practice to guarantee that the selection is truly random across the full spectrum of individual wealth.

The statistics for individual wealth pertain to the United Kingdom in the period 2014–2016. Clearly there will be differences in detail between different nations and they will change over time.

Table 2
Summary of sample sizes.

Description	Actual sample size	Under 12% rule		
		Degenerate, lower limit on sample size	Central figure for minimum sample size	Range of minimum sample sizes as multiplier deterministic component, r_0 , varies
Multiplier Model of VPF survey	–	1860 ^a	2104 ($r_0 = r_{0\max}/2$)	1860–2835
Utility Model of VPF survey	–	1860 ^b	2648	–
Extended Utility Model of VPF survey	–	1860 ^c	2958 ($r_0 = r_{0\max}/2$)	2648–3887
UK VPF survey by Chilton et al.	167	–	–	–

^a Deterministic multiplier: $r_0 = r_{0\max}$.

^b Risk-aversion = zero.

^c Deterministic multiplier: $r_0 = r_{0\max}$ and risk-aversion = zero.

Nevertheless the numbers found for minimum sample size are likely to give good guidance to those intending to undertake a survey measurement of any parameter that is strongly wealth dependent in a developed country.

The measurement by survey of household wealth, as carried out by the Office of National Statistics, enjoys a significant advantage over the estimation of a general, non-market parameter, in that it is possible to break down a household's wealth into assets for which a market exists: bank accounts, shares, property etc. This will allow a reasonably accurate quantification of the monetary value of each class of holding, and, as long as all assets are identified, will permit a similarly accurate and testable quantification of total wealth. Unfortunately such a methodical mode of proceeding is not available to those attempting a stated preference survey, for example a survey estimation of the VPF. In such a case, while choosing a sample truly randomly and ensuring it is of adequate size can provide the desired β -level precision for the survey, the accuracy of the resulting figure cannot be relied upon for the reasons outlined by Fujiwara and Campbell [30], as discussed in Section 1.2 above.

6. Conclusions

The large variation in individual wealth in a developed country like the UK means that the sample size for the survey measurement of a continuous parameter that is strongly dependent on wealth will always need to be at least ~75% larger than the sample size of a political opinion poll giving the equivalent level of precision. This sets a degenerate lower limit of about 2000 on the minimum sample size for the survey measurement of a such a parameter at an achieved precision equivalent to the 3% margin of error common in political opinion polls.

In the case of the UK “value of a prevented fatality”, the sample size needs to be about 3000, based on the Extended Utility Model likely to be most realistic, if the precision of the survey is to match that of the political opinion polls that people are used to seeing reported in their national media.

It should be emphasised that the selection of the sample needs to be fully random across the whole population for the survey to have validity. Particular care is needed if the opinions of wealthy people are to be captured adequately.

It is clear that small scale opinion surveys cannot provide adequate precision in the measurement of continuous parameters, such as the VPF, that have a strong dependence on wealth.

No confidence can be placed in the figure for the VPF used as a safety yardstick by the UK government because the sample size used in its foundational opinion survey, 167, is less than a tenth of what would be required to give reasonable precision. This adds

to the catalogue of well documented problems besetting the UK VPF, the use of which is not justified.

The Fujiwara and Campbell caveat needs always to be borne in mind, of course, that stated preferences do not necessarily represent true preferences. Hence even when a survey has included a wide enough spread of randomly chosen respondents to provide good precision, the test of accuracy may still not be passed.

Declaration of Competing Interest

The author declares the following financial interests/personal relationships which may be considered as potential competing interests: the author is employed by the University of Bristol as Professor of Risk Management; he is a director of Michaelmas Consulting Ltd, which has sponsored open-access publication of this article.

Acknowledgements

The author would like to thank Dr Ian Waddington (Analytic Eye Ltd.), Dr John May (University of Bristol), Prof. Richard Taylor (University of Bristol) and Dr Bill Bardo for their helpful advice and comments.

Appendix A. Deriving the mean and variance from tabular data for cumulative probability

The probability distribution for a general random variable, X , is sometimes given in the form of a table of values, x_i , for points at regular intervals of cumulative probability, $F_X(x_i)$. In general, the table will contain $n + 1$ pairs, $(x_i, F_X(x_i))$, $i = 0, 1, 2, \dots, n$, and the n intervals of $x, x_{i-1} \rightarrow x_i$, $i = 1, 2, 3, \dots, n$, will have a length, $x_i - x_{i-1}$, that can be expected to vary. By contrast, the regularity of the intervals of F_X means that they will all have the same length:

$$\Delta F_X = \frac{1}{n} \quad (\text{A.1})$$

Here, for example, $n = 4$ if the data are given in quartiles, $n = 10$ if given in deciles and $n = 100$ for a table of percentiles such as Table 1.

The probability density function, $f_X(x)$, may be calculated for each interval of x after assuming a piece-wise uniform increase in cumulative probability, $F_X(x)$, between quantiles. Thus

$$f_X(x) = \frac{F_X(x_i) - F_X(x_{i-1})}{x_i - x_{i-1}} = \frac{\Delta F_X}{x_i - x_{i-1}} \quad \text{for } x_{i-1} \leq x < x_i \quad (\text{A.2})$$

$$= f_{X,i} \quad i = 1, 2, \dots, n$$

in which 0 is the index of the first data point, $(x_0, F_X(x_0))$ and n is the index of the last data point, $(x_n, F_X(x_n))$, where $F_X(x_0) = 0$ and $F_X(x_n) = 1$. The expected value of X is then:

$$\begin{aligned}
E(X) &= \int_{x=x_0}^{x_n} x f_X(x) dx = \int_{x_0}^{x_1} x f_{X,1} dx + \int_{x_1}^{x_2} x f_{X,2} dx + \dots \\
&+ \int_{x_{i-1}}^{x_i} x f_{X,i} dx + \dots + \int_{x_{n-1}}^{x_n} x f_{X,n} dx = f_{X,1} \int_{x_0}^{x_1} x dx \\
&+ f_{X,2} \int_{x_1}^{x_2} x dx + \dots + f_{X,i} \int_{x_{i-1}}^{x_i} x dx + \dots + f_{X,n} \int_{x_{n-1}}^{x_n} x dx \\
&= f_{X,1} \frac{x_1^2 - x_0^2}{2} + f_{X,2} \frac{x_2^2 - x_1^2}{2} + \dots \\
&+ f_{X,i} \frac{x_i^2 - x_{i-1}^2}{2} + \dots + f_{X,n} \frac{x_n^2 - x_{n-1}^2}{2}
\end{aligned} \quad (A.3)$$

Expanding the differences between squares and using Eq. (A.2), allows Eq. (A.3) to be transformed into:

$$\begin{aligned}
E(X) &= \frac{\Delta F_X}{x_1 - x_0} \frac{(x_1 - x_0)(x_1 + x_0)}{2} + \frac{\Delta F_X}{x_2 - x_1} \frac{(x_2 - x_1)(x_2 + x_1)}{2} \\
&+ \dots + \frac{\Delta F_X}{x_i - x_{i-1}} \frac{(x_i - x_{i-1})(x_i + x_{i-1})}{2} + \dots \\
&+ \frac{\Delta F_X}{x_n - x_{n-1}} \frac{(x_n - x_{n-1})(x_n + x_{n-1})}{2} \\
&= \frac{\Delta F_X}{2} ((x_0 + x_1) + (x_1 + x_2) + \dots + (x_{i-1} + x_i) + \dots + (x_{n-1} + x_n))
\end{aligned} \quad (A.4)$$

where the last line follows by cancellation of the factor, $x_i - x_{i-1}$, $i = 1, 2, \dots, n$.

Using Eq. (A.1), the expected or mean value may thus be restated as

$$E(X) = \frac{1}{2n} \left(x_0 + 2 \sum_{i=1}^{n-1} x_i + x_n \right) \quad (A.5)$$

Meanwhile the second moment of X is given by:

$$\begin{aligned}
E(X^2) &= \int_{x=x_0}^{x_n} x^2 f_X(x) dx = \int_{x_0}^{x_1} x^2 f_{X,1} dx + \int_{x_1}^{x_2} x^2 f_{X,2} dx + \dots \\
&+ \int_{x_{i-1}}^{x_i} x^2 f_{X,i} dx + \dots + \int_{x_{n-1}}^{x_n} x^2 f_{X,n} dx = f_{X,1} \int_{x_0}^{x_1} x^2 dx \\
&+ f_{X,2} \int_{x_1}^{x_2} x^2 dx + \dots + f_{X,i} \int_{x_{i-1}}^{x_i} x^2 dx + \dots + f_{X,n} \int_{x_{n-1}}^{x_n} x^2 dx \\
&= f_{X,1} \frac{x_1^3 - x_0^3}{3} + f_{X,2} \frac{x_2^3 - x_1^3}{3} + \dots + f_{X,i} \frac{x_i^3 - x_{i-1}^3}{3} + \dots \\
&+ f_{X,n} \frac{x_n^3 - x_{n-1}^3}{3} = f_{X,1} (x_1 - x_0) \frac{x_1^2 + x_1 x_0 + x_0^2}{3} \\
&+ f_{X,2} (x_2 - x_1) \frac{x_2^2 + x_2 x_1 + x_1^2}{3} \dots \\
&+ f_{X,i} (x_i - x_{i-1}) \frac{x_i^2 + x_i x_{i-1} + x_{i-1}^2}{3} + \dots \\
&+ f_{X,n} (x_n - x_{n-1}) \frac{x_n^2 + x_n x_{n-1} + x_{n-1}^2}{3}
\end{aligned} \quad (A.6)$$

Substituting from Eq. (A.2) into Eq. (A.6) gives:

$$E(X^2) = \frac{\Delta F_X}{3} \left(x_1^2 + x_1 x_0 + x_0^2 + x_2^2 + x_2 x_1 + x_1^2 + \dots + x_i^2 + x_i x_{i-1} + x_{i-1}^2 + \dots + x_n^2 + x_n x_{n-1} + x_{n-1}^2 \right) \quad (A.7)$$

After grouping values and using Eq. (A.1), this gives:

$$E(X^2) = \frac{1}{3n} \left(x_0^2 + 2 \sum_{i=1}^{n-1} x_i^2 + x_n^2 + \sum_{i=1}^n x_i x_{i-1} \right) \quad (A.8)$$

The variance of X may be calculated by substituting Eqs. (A.5) and (A.8) into:

$$\text{var}(X) = E(X^2) - (E(X))^2 \quad (A.9)$$

Appendix B. The confidence interval, the margin of error used in political opinion polls and β -level precision

B.1. Modelling an opinion poll as a binomial random process

A binomial random process is a natural model to use when interpreting opinion polls held in advance of the formal vote they are intended to predict. X is then a random variable representing the number of people found to hold a certain view in a poll of n people.

The binomial process is linked to the normal distribution by the De Moivre-Laplace Limit Theorem [53], which states that, for any numbers, a and b , on the real axis:

$$\lim_{n \rightarrow \infty} P \left(a \leq \frac{\frac{X}{n} - p}{\sqrt{p(1-p)}/\sqrt{n}} \leq b \right) = \Phi(b) - \Phi(a) \quad (B.1)$$

Here p is the fraction holding the specified opinion in the population as a whole and

$$\Phi(x) = \int_{-\infty}^x f_Z(z) dz \quad (B.2)$$

is the cumulative probability to x of a standard normal distribution, for which the probability density is

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \quad (B.3)$$

As a matter of practicality, of course, no opinion poll can consult the infinite number of people required by Eq. (B.1). But fortunately the distribution will retain approximate normality for finite but large n , when it is possible to write:

$$P(a \leq G(X, n, p) \leq b) \approx \Phi(b) - \Phi(a) \quad (B.4)$$

where the random number, $G(X, n, p)$, is given by:

$$G(X, n, p) = \frac{\frac{X}{n} - p}{\frac{1}{\sqrt{n}} \sqrt{p(1-p)}} \quad (B.5)$$

It is useful to know that the expected value of $G(X, n, p)$ will be zero, a result that follows from applying the expectation operator, $E(\cdot)$, to both sides of Eq. (B.5) and taking account of the fact that n and p are both deterministic:

$$E(G(X, n, p)) = \frac{1}{\frac{1}{\sqrt{n}} \sqrt{p(1-p)}} \left(\frac{1}{n} E(X) - p \right) = 0 \quad (B.6)$$

where the last step follows from the fact that for a binomial distribution, $E(X) = np$.

B.2. The confidence interval

This subsection will calculate the minimum length of the random interval that has a specified probability, β , of containing the true population fraction, p . This minimum length is carried over to the confidence interval that is quoted for the sample estimate of that fraction.

The interval, (a, b) , connecting any positions, a and $b : b > a$, on the real axis will have length, $\omega = b - a$. If the positions, a and b , are allowed to vary while keeping the length, ω , of the interval constant, then the probability of containing G in the interval between a and b , $P(a \leq G \leq b)$, will be greatest when $a = -b$ and the interval, $(-b, b)$, is symmetrical about the origin. As $E(G) = 0$ by Eq. (B.6), the resulting interval will also be centred on $E(G)$. Moreover, if the probability, $P(a \leq G \leq b)$, of capturing G in the interval between a and b is specified, and the length, $\omega = b - a$, of that interval is varied to achieve the required probability, then this length will be minimised when that interval straddles the ori-

gin symmetrically. A demonstration of these results is given in [Appendix C](#).

Putting $a = -b$ to give the shortest interval that will satisfy the specified probability and combining Eqs. (B.4), (B.5) and (C.3) from [Appendix C](#) gives:

$$P\left(-b \leq \frac{\frac{X}{n} - p}{\frac{1}{\sqrt{n}}\sqrt{p(1-p)}} \leq b\right) \approx \Phi(b) - \Phi(-b) = 2\Phi(b) - 1 \quad (\text{B.7})$$

Rearranging the inequality contained within the probability brackets gives

$$P\left(\frac{X}{n} - b \frac{1}{\sqrt{n}}\sqrt{p(1-p)} \leq p \leq \frac{X}{n} + b \frac{1}{\sqrt{n}}\sqrt{p(1-p)}\right) \approx 2\Phi(b) - 1 \quad (\text{B.8})$$

For example, when $b = 1.96$, $\Phi(b) = 0.975$ and so $2\Phi(b) - 1 = 0.95$ or 95% when the probability is expressed as a percentage. Thus the probability is 95% that the population fraction will be contained within the random interval defined in Eq. (B.8) with $b = 1.96$. The length of this interval is

$$w = \frac{2 \times b}{\sqrt{n}}\sqrt{p(1-p)} = \frac{2 \times 1.96}{\sqrt{n}}\sqrt{p(1-p)} \quad (\text{B.9})$$

which is the shortest length able to contain the true population fraction, p , with 95% probability. The interval is centred on the random variable, X/n , which has an expected value, $E(X)/n = p$, because X is binomially distributed.

When the opinion poll has been completed and the number of people with a specified opinion has been found to be x , then replacing the random variable, X , with the outturn, x , in the main bracketed term defines the $\beta\%$ confidence interval:

$$\left(\frac{x}{n} - b^{(\beta)} \frac{1}{\sqrt{n}}\sqrt{p(1-p)}, \frac{x}{n} + b^{(\beta)} \frac{1}{\sqrt{n}}\sqrt{p(1-p)}\right) \quad (\text{B.10})$$

where $b^{(\beta)}$ is the solution of the implicit equation:

$$\Phi(b) = \frac{1}{2} \left(1 + \frac{\beta}{100}\right) \quad (\text{B.11})$$

The length of the confidence interval, now named w_{poll} to tie it explicitly to the political opinion poll, is:

$$w_{\text{poll}}^{(\beta)} = \frac{2b^{(\beta)}}{\sqrt{n}}\sqrt{p(1-p)} \quad (\text{B.12})$$

This equation makes it explicit that the length of the confidence interval will depend on the probability, β , associated with b (the two quantities are linked via Eq. (B.11)).

As the confidence interval and the probability interval share the same length, it follows that the length of the confidence interval specifies the shortest interval length with the potential to contain p with probability, $\beta\%$, in a similar survey to be carried out in the future.

B.3. The margin of error and β -level precision

A plebiscite normally offers two, mutually exclusive options, as was the case, for example, in the EU Referendum held in the United Kingdom on 23 June 2016. Here the electorate was invited to choose between two options: either for the UK to remain in the European Union or for the country to leave that economic and political grouping, with the Government undertaking to accept the people's majority verdict. The result was close, but in the event, 16,141,241 (48.1%) people voted to remain and 17,410,742 people (51.9%) voted to leave [54].

The objective of the numerous opinion polls published in national newspapers in advance of the referendum was to estimate

the actual fraction, p , of those who would vote to leave the EU, should the ballot be held on the day the opinion survey was performed. From Eq. (B.12), the length of the confidence interval will depend only on the number, n , of people questioned if p is known (perhaps known approximately, between limits). For example, when $p = 0.5$, then

$$w_{\text{poll}}^{(\beta)} = \frac{b^{(\beta)}}{\sqrt{n}} \quad (\text{B.13})$$

This equation allows the length of the confidence interval to be calculated as soon as the probability, β , and the number of people in the survey, n , are known. Alternatively, specification of the probability, β , and the desired length, $w_{\text{poll}}^{(\beta)}$, of the confidence interval will allow a determination of the necessary sample size, n , in advance of the survey being carried out.

If p is not known, but is estimated as $p \approx x/n$ after the survey has been carried out, then (B.12) becomes:

$$w_{\text{poll}}^{(\beta)} = \frac{2b^{(\beta)}}{\sqrt{n}}\sqrt{\frac{x}{n}\left(1 - \frac{x}{n}\right)} \quad (\text{B.14})$$

Since Eq. (B.14) relies on the result of the survey, it cannot be solved in advance. However, the dependence on x/n in Eq. (B.14) may be removed in two ways that lead to similar results. First, it may be noted that, in general, the expression $y(1-y)$ experiences a maximum of 0.25 at $y = 0.5$. This prompts the consideration of $w_{\text{pollmax}} = \max_{x/n} w_{\text{poll}}^{(\beta)}$, as a conservative (i.e. overlarge) estimate of $w_{\text{poll}}^{(\beta)}$. This implies substituting $x/n = 0.5$ in the right-hand side of Eq. (B.14), which then produces:

$$w_{\text{pollmax}}^{(\beta)} = \frac{b^{(\beta)}}{\sqrt{n}} \quad (\text{B.15})$$

But another way of proceeding is to assume that the contest is going to be close, as in the case of the EU Referendum, implying $x/n \approx 0.5$. This leads directly to Eq. (B.13).

While the exact nature of the derivation of Eq. (B.15) is mathematically satisfying, its usefulness in practice rests on the *de facto* approximation

$$w_{\text{poll}}^{(\beta)} \approx w_{\text{pollmax}}^{(\beta)} \quad (\text{B.16})$$

which returns attention to Eq. (B.13). Despite being approximate when $x/n \neq 0.5$, this equation is significantly robust against variations in x/n , as shown by Fig. 10. For example, the error introduced by applying Eq. (B.13) rather than the full form of Eq. (B.14) is less than 10% over the range: $0.283 \leq x/n \leq 0.717$.

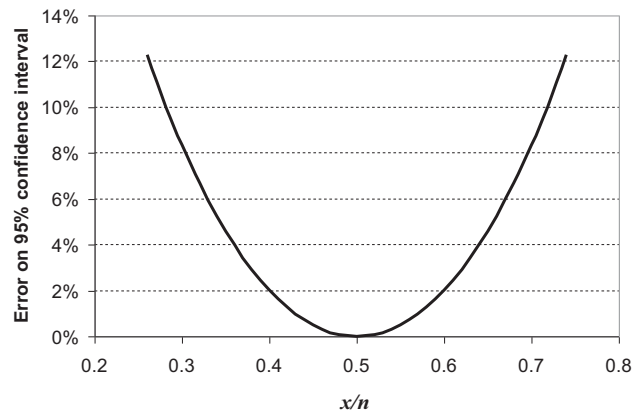


Fig. 10. Error on 95% confidence interval against assumed value of x/n .

Opinion poll results are normally quoted as $x/n \pm m$, where m is the margin of error, defined as:

$$m^{(\beta)} = \frac{w_{poll}^{(\beta)}}{2} \quad (\text{B.17})$$

The bracketed superscripts in Eq. (B.17) make it clear that the margin of error will depend, as does $w_{poll}^{(\beta)}$, on the selection of the probability, β , at which b is found (Eq. (B.11)).

Combining Eqs. (B.13) and (B.17) gives:

$$n = \left(\frac{b^{(\beta)}}{2m^{(\beta)}} \right)^2 \quad (\text{B.18})$$

While the margin of error, $m^{(\beta)}$, is an absolute figure, the expression, $m^{(\beta)}/(x/n)$, gives the tolerance relative to the estimated fraction, x/n . From Eqs. (B.13) and (B.18)

$$\frac{m^{(\beta)}}{x/n} \Big|_{x/n=0.5} = \frac{m^{(\beta)}}{0.5} = w_{poll}^{(\beta)} = \frac{b^{(\beta)}}{\sqrt{n}} \quad (\text{B.19})$$

The expression, $b^{(\beta)}/\sqrt{n}$, gives a measure, conditioned on β , of the precision of the prediction for the later referendum or election. Similarly the relative length, $\lambda_{poll}^{(\beta)}$, of the $\beta\%$ confidence interval, relative to x/n when $x/n = 0.5$, which will be double $b^{(\beta)}/\sqrt{n}$, must also be a measure of the precision of the prediction. This new term, $\lambda_{poll}^{(\beta)}$, is again conditioned on the probability and hence confidence level, β , chosen, and it will be termed the “ β -level precision”:

$$\lambda_{poll}^{(\beta)} = \frac{w_{poll}^{(\beta)}}{x/n} \Big|_{x/n=0.5} = 2 \frac{m^{(\beta)}}{x/n} \Big|_{x/n=0.5} = \frac{2b^{(\beta)}}{\sqrt{n}} \quad (\text{B.20})$$

UK political opinion polls normally use a 95% confidence interval so that $\beta = 95\%$, which implies from Eq. (B.11) that $b = 1.96$. The margin of error is conventionally set equal to 3% in the UK, so that $m^{(0.95)} = 0.03$. The equivalent β -level precision is 4 times this figure:

$$\lambda_{poll}^{(0.95)} = 4m^{(0.95)} = 0.12 \quad (\text{B.21})$$

Appendix C. Minimising the interval containing the population average with a specified probability in an opinion poll

C.1. Maximising the probability of an interval of given length containing the population fraction, p

Consider a general, symmetrical interval, (a, b) , as shown in Fig. 11(a)–11(c), where

$$a = -b \quad (\text{C.1})$$

By the symmetry evident in Fig. 11, of the probability density of the standard normal distribution, if any number, c , is chosen, then the cumulative probability to $-c$ is the complement of the cumulative probability to c :

$$\Phi(-c) = 1 - \Phi(c) \quad (\text{C.2})$$

Using Eqs. (C.1) and (C.2) in Eq. (B.4) produces

$$P(a \leq G(X, n, p) \leq b) \approx 2\Phi(b) - 1 \quad (\text{C.3})$$

Now consider the case where b is reduced by a difference, $d : d > 0$, to $b - d$. To maintain the length of the interval at $b - a$, it is necessary to reduce a by the same amount to $a - d$. The probability associated with this new interval, $(a - d, b - d)$, is given, using Eq. (B.4), as:

$$P(a - d \leq G(X, n, p) \leq b - d) \approx \Phi(b - d) - \Phi(a - d) \quad (\text{C.4})$$

But, using Eq. (C.1) and then Eq. (C.2),

$$\Phi(a - d) = \Phi(-b - d) = 1 - \Phi(b + d) \quad (\text{C.5})$$

Hence, on substituting Eq. (C.5) into Eq. (C.4)

$$P(a - d \leq G(X, n, p) \leq b - d) \approx \Phi(b - d) + \Phi(b + d) - 1 \quad (\text{C.6})$$

The change in probability, $\Delta P(-d)$, caused by reducing b by an amount, d , while maintaining the length of the interval constant may be found by subtracting Eq. (C.3) from Eq. (C.6):

$$\begin{aligned} \Delta P(-d) &\approx \Phi(b - d) + \Phi(b + d) - 1 - (2\Phi(b) - 1) \\ &= \Phi(b + d) - \Phi(b) - \{\Phi(b) - \Phi(b - d)\} \end{aligned} \quad (\text{C.7})$$

Eq. (C.7) remains valid when d is added to b (when $a \rightarrow a + d$ to maintain a constant interval length). This leads to an increase in the area under the probability density curve of $\Phi(b + d) - \Phi(b)$ on the right but a decrease of $\Phi(a + d) - \Phi(a)$ on the left. But

$$\begin{aligned} \Phi(a + d) - \Phi(a) &= \Phi(-b + d) - \Phi(-b) \\ &= 1 - \Phi(b - d) - (1 - \Phi(b)) = \Phi(b) - \Phi(b - d) \end{aligned} \quad (\text{C.8})$$

It follows that

$$\Delta P(+d) = \{\Phi(b + d) - \Phi(b)\} - \{\Phi(b) - \Phi(b - d)\} \quad (\text{C.9})$$

so that, on comparing Eqs. (C.7) and (C.9), $\Delta P(+d) = \Delta P(-d)$, as claimed.

To proceed further it is now necessary to distinguish between the three exclusive and exhaustive possibilities: $d \leq b$, $b < d \leq 2b$ and $d > 2b$. See Figs. 11(a), 11(b) and 11(c).

$d \leq b$ Referring to Fig. 11(a) Eq. (C.7) may be expanded using the definition of $\Phi(x)$ in Eq. (B.2) and using the following three intermediate points in the various integrations: $b - d$, b and $b + d$:

$$\begin{aligned} \Delta P(-d) &\approx \left[\int_{-\infty}^b f_Z(z) dz + \int_b^{b+d} f_Z(z) dz \right] - \int_{-\infty}^b f_Z(z) dz \\ &\quad - \left\{ \left[\int_{-\infty}^{b-d} f_Z(z) dz + \int_{b-d}^b f_Z(z) dz \right] - \int_{-\infty}^{b-d} f_Z(z) dz \right\} \end{aligned} \quad (\text{C.10})$$

Hence

$$\Delta P(-d) \approx \int_b^{b+d} f_Z(z) dz - \int_{b-d}^b f_Z(z) dz \quad (\text{C.11})$$

The terms on the right-hand side of Eq. (C.11) are the shaded areas shown on the right-hand side of Fig. 11 (a). It is visually clear that the first term in Eq. (C.11) is smaller than the second, indicating that $\Delta P(-d)$ must be negative: $\Delta P(-d) < 0$.

More formally, by the definition of the standard normal probability density (Eq. (B.3)), it is plain that the density will be higher the closer z is to 0:

$$f_Z(|z_1|) > f_Z(|z_2|) \text{ for } |z_1| < |z_2| \quad (\text{C.12})$$

Since in this case, $0 < d \leq b$, it follows that

$$f_Z(z) > f_Z(b) \text{ for } b - d < z < b \quad (\text{C.13})$$

While

$$f_Z(z) < f_Z(b) \text{ for } b < z < b + d \quad (\text{C.14})$$

Thus the second integrand in Eq. (C.11) will always be larger than the first and, since the length of the integration interval is the same ($= d$), it follows that

$$\int_b^{b+d} f_Z(z) dz < \int_{b-d}^b f_Z(z) dz \quad (\text{C.15})$$

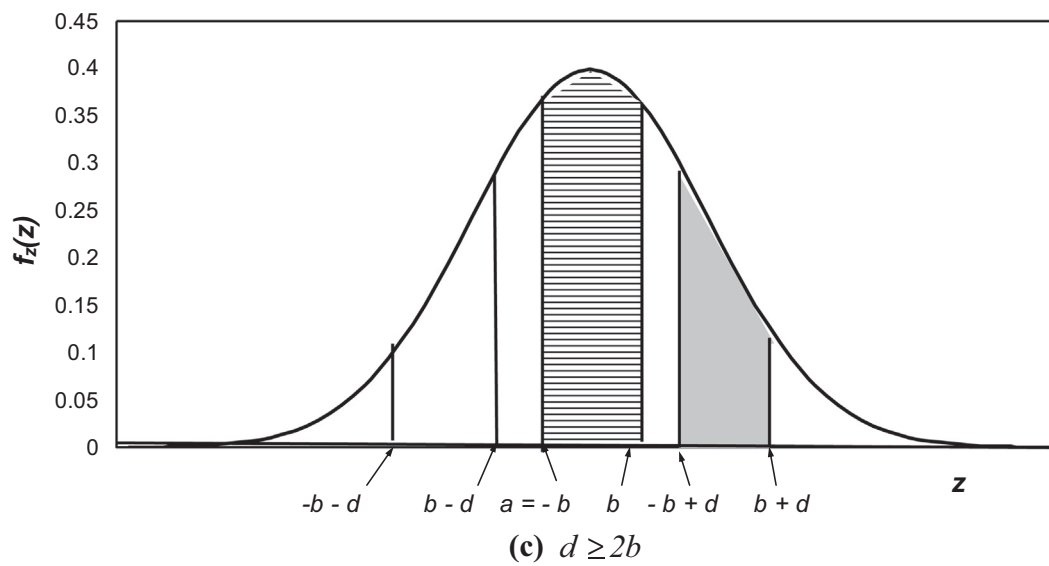
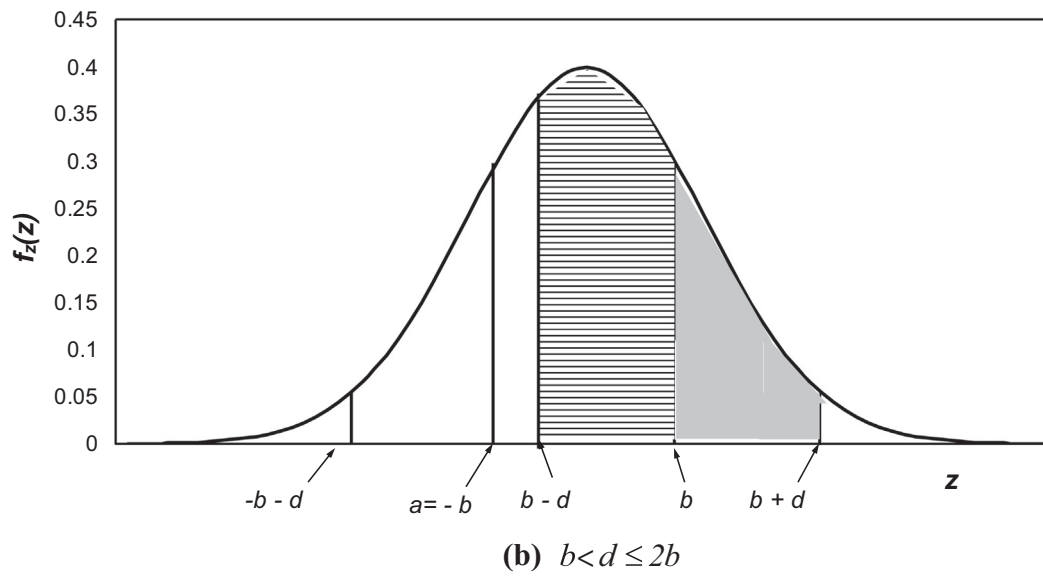
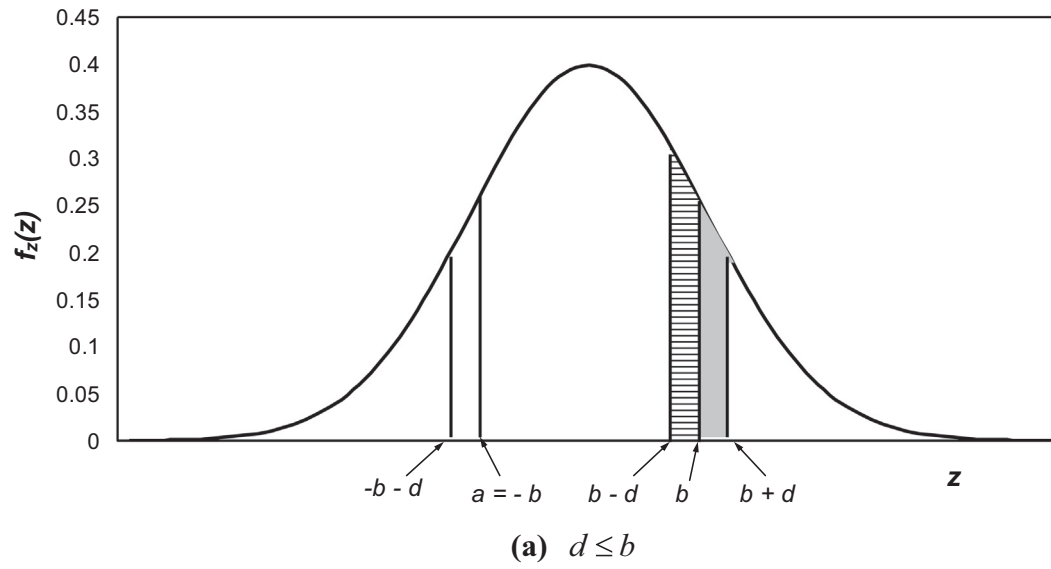


Fig. 11. Probability density for the standard normal distribution.

Hence

$$\Delta P(-d) < 0 \quad (\text{C.16})$$

$b < d \leq 2b$ This case is illustrated in Fig. 11(b), which differs from Fig. 11(a) in that the curve defining the area, $\int_{b-d}^b f_Z(z) dz$, undergoes a maximum between the two integration limits. However, the areas under the probability density curve in Eq. (C.11) remain contiguous, and the rationale remains the same as for the case where $d < b$. As before, $f_Z(z) > f_Z(b)$ for $b - d < z < b$, including for the limiting case in this category, where $d = 2b$.

$d > 2b$ Fig. 11(c) shows this case. In expanding Eq. (C.7), 5 intermediate points are now used as limits in the various integrations: $b - d, -b, b, -b + d$ and $b + d$. Hence

$$\begin{aligned} \Delta P(-d) \approx & \left[\int_{-\infty}^{b-d} f_Z(z) dz + \int_{b-d}^{-b} f_Z(z) dz + \int_{-b}^b f_Z(z) dz + \int_b^{-b+d} f_Z(z) dz + \int_{-b+d}^{b+d} f_Z(z) dz \right] \\ & - \left[\int_{-\infty}^{b-d} f_Z(z) dz + \int_{b-d}^{-b} f_Z(z) dz + \int_{-b}^b f_Z(z) dz \right] \\ & - \left\{ \left[\int_{-\infty}^{b-d} f_Z(z) dz + \int_{b-d}^{-b} f_Z(z) dz + \int_{-b}^b f_Z(z) dz \right] - \int_{-\infty}^{b-d} f_Z(z) dz \right\} \\ & = \int_b^{-b+d} f_Z(z) dz + \int_{-b+d}^{b+d} f_Z(z) dz - \int_{b-d}^{-b} f_Z(z) dz - \int_{-b}^b f_Z(z) dz \end{aligned} \quad (\text{C.17})$$

But, by Eq. (C.2)

$$\begin{aligned} \int_b^{-b+d} f_Z(z) dz &= \Phi(-b+d) - \Phi(b) = 1 - \Phi(b-d) - (1 - \Phi(-b)) \\ &= \Phi(-b) - \Phi(b-d) = \int_{b-d}^{-b} f_Z(z) dz \end{aligned} \quad (\text{C.18})$$

Hence the 1st and 3rd terms in the final line of Eq. (C.17) cancel and:

$$\Delta P(-d) \approx \int_{-b+d}^{b+d} f_Z(z) dz - \int_{-b}^b f_Z(z) dz \quad (\text{C.19})$$

These two areas are shaded in Fig. 11(c), from which it is visually clear that the second area is bigger than the first. But proceeding formally, since now $d > 2b$, it follows that $-b + d > b$. Hence by condition (C.12)

$$f_Z(z) < f_Z(b) \text{ for } -b + d < z < b + d \quad (\text{C.20})$$

By contrast

$$f_Z(z) > f_Z(b) \text{ for } -b < z < b \quad (\text{C.21})$$

Since the length of the integration interval is the same in the two integrals ($=2b$), it follows that the second integral will be larger than the first and so, once again condition (C.16) applies: $\Delta P(-d) < 0$.

The necessary result is thus demonstrated: if the positions, a and b , are allowed to vary while keeping the length, $\omega = b - a$, of the interval constant, then the probability of containing G in the interval between a and b , $P(a \leq G \leq b)$, will be greatest when $a = -b$ and $\omega = 2b$. Any deviation from symmetry about the origin will cause the probability to fall.

An easily understood, partial corroboration of the finding comes from considering an interval confined to the negative real axis. The symmetry about the origin of the probability density, $f_Z(z)$, means that the maximum probability such an interval can generate is 50%. A similar conclusion is reached when the interval is limited to the positive real axis. It is thus immediately clear that the probabilities normally of interest, 90%, 95% or 99%, can be produced only by an interval that straddles the origin. The fuller analysis above shows that the maximum probability will occur when the interval not only straddles the origin but is symmetrical about it.

C.2. Minimising the length of the interval containing, with any specified probability, the population fraction, p

Consider a general interval, (h, k) , of length, $k - h$, which has a probability, $\theta_{h,k}$, of including $G(X, n, p)$: $P(h \leq G(X, n, p) \leq k) = \theta_{h,k}$. By the result of Section C.1, that probability will be maximised for intervals of the same length, $k - h$, when the interval straddles the origin symmetrically, in which case the interval is $(-\kappa, \kappa)$, where

$$\kappa = \frac{k - h}{2} \quad (\text{C.22})$$

The interval length is now 2κ . From Eqs. (B.4) and (B.2)

$$P(-\kappa \leq G(X, n, p) \leq \kappa) = \theta_{-\kappa, \kappa} = \int_{-\kappa}^{\kappa} f_Z(z) dz \quad (\text{C.23})$$

This probability may be compared with the case considered in the previous Section, where the interval was $(-b, b)$, of length $2b$:

$$P(-b \leq G(X, n, p) \leq b) = \theta_{-b, b} = \int_{-b}^b f_Z(z) dz \quad (\text{C.24})$$

If $\kappa < b$, Eq. (C.24) may be written:

$$\begin{aligned} \theta_{-b, b} &= \int_{-b}^{-\kappa} f_Z(z) dz + \int_{-\kappa}^{\kappa} f_Z(z) dz + \int_{\kappa}^b f_Z(z) dz \\ &= \theta_{-\kappa, \kappa} + \int_{-b}^{-\kappa} f_Z(z) dz + \int_{\kappa}^b f_Z(z) dz \end{aligned} \quad (\text{C.25})$$

Since $f_Z(z) > 0$ for all z and, in particular, for $-b \leq z \leq -\kappa$ and $\kappa \leq z \leq b$, it follows that

$$\theta_{-b, b} > \theta_{-\kappa, \kappa} \text{ if } \kappa < b \quad (\text{C.26})$$

Furthermore, the probability associated with an interval that located symmetrically about the origin will be greater than that produced by any other interval of the same length:

$$\theta_{-\kappa, \kappa} > \theta_{h, k} \text{ for } (h, k) \neq (-\kappa, \kappa) \quad (\text{C.27})$$

It follows from the conjunction of conditions (C.26) and (C.27) that:

$$\theta_{-b, b} > \theta_{h, k} \text{ when } k - h < 2b \quad (\text{C.28})$$

Condition (C.28) shows that the probability generated by an interval that is symmetrical about the origin is greater than that associated with any interval, wherever it occurs on the real axis, if that second interval is shorter.

Hence for any specified probability of containing $G(X, n, p)$, the length of the interval will be minimised by selecting the probability-compliant interval that straddles the origin symmetrically. By the same token, the interval will straddle $E(G)$ symmetrically, as $E(G) = 0$ by Eq. (B.6).

Appendix D. Confidence interval for a general population average

Referring to Appendix B, the random number, X/n , may be regarded as the average, \bar{Y} , of n independent trials of a Bernoulli process in each of which the random variable, Y , takes the value, 1, with probability, p , and 0 with probability, $1 - p$. Meanwhile the true mean of the Bernoulli process is given by

$$\mu = E(Y) = p \times 1 + (1 - p) \times 0 = p \quad (\text{D.1})$$

and the standard deviation by:

$$\begin{aligned} \sigma &= \sqrt{\text{var}(Y)} = \sqrt{E(Y^2) - (E(Y))^2} \\ &= \sqrt{p \times 1^2 + (1 - p) \times 0^2 - p^2} = \sqrt{p(1 - p)} \end{aligned} \quad (\text{D.2})$$

Eq. (B.1) may thus be recast as:

$$\lim_{n \rightarrow \infty} P\left(a \leq \frac{\bar{Y} - \mu}{\sigma/\sqrt{n}} \leq b\right) = \Phi(b) - \Phi(a) \quad (\text{D.3})$$

which is a statement of the Central Limit Theorem [53], which has general application to averages from a wide variety of distributions, including those applying to a continuous parameter.

Taking the case of a continuous parameter, putting $a = -b$ to give the shortest interval and allowing for a finite but large number, n , of samples gives the approximate equality:

$$P\left(-b \leq \frac{\bar{Y} - \mu}{\sigma/\sqrt{n}} \leq b\right) \approx 2\Phi(b) - 1 \quad (\text{D.4})$$

so that

$$P\left(\bar{Y} - b \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{Y} + b \frac{\sigma}{\sqrt{n}}\right) \approx 2\Phi(b) - 1 \quad (\text{D.5})$$

For example $b = 1.96$ gives

$$P\left(\bar{Y} - 1.96 \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{Y} + 1.96 \frac{\sigma}{\sqrt{n}}\right) \approx 0.95 \quad (\text{D.6})$$

By the analysis of Appendix C, the length of the interval cited in Eq. (D.5), $2b^{(\beta)} \frac{\sigma}{\sqrt{n}}$, is the shortest that can offer a $\beta\%$ probability of containing the population average, μ . To fulfil this probability, the interval needs to be centred on \bar{Y} , the expected value of which is given by: $E(\bar{Y}) = \mu$.

The $\beta\%$ confidence interval is centred on the actual sample average, \bar{y} :

$$\left(\bar{y} - b^{(\beta)} \frac{\sigma}{\sqrt{n}}, \bar{y} + b^{(\beta)} \frac{\sigma}{\sqrt{n}}\right) \quad (\text{D.7})$$

when the standard deviation, σ , is known. This interval is a realisation of the random interval set down in Eq. (D.5).

Each of the terms in expression (D.7) is deterministic, a reflection of the fact that the uncertainty has disappeared, so that the interval is deterministic and no longer random. Therefore the probability operator, $P(\cdot)$, ceases to have meaning and it cannot be claimed that the interval described in expression (D.7) has a probability, β , of containing the population average, μ . However, reference may be made to the provenance of the deterministic interval, namely that it resulted from a process that obeyed Eq. (D.5), that is to say a process that had a probability, β , of capturing the mean within the interval. The term, “confidence”, can be ascribed to this prior probability. Furthermore, uncertainty could be re-introduced into the overall system of measurement by repeating the estimation process many times using different samples, and then allowing the measurer to choose at random just one of the resultant different realisations of the interval. A fraction, $\sim \beta$, of the intervals so produced, will contain the true mean, μ , so that the probability of choosing an interval containing the true mean will be $\sim \beta$ (an infinite number of realisations would be needed to render the probability exactly equal to β). The sample average, as seen by the selector, will now be random again, so that Eq. (D.5) applies once more. The term, “confidence”, can also be ascribed to this “posterior” probability.

The length of the $\beta\%$ confidence interval can be found by subtracting the first from the second term in expression (D.7):

$$w = 2b^{(\beta)} \frac{\sigma}{\sqrt{n}} \quad (\text{D.8})$$

This is clearly the same length as the probability interval given in Eq. (D.5), and is thus the shortest length, measured in the units of the average value, μ , that can offer a $\beta\%$ probability of containing μ .

The relative length, $\lambda^{(\beta)}$, of the $\beta\%$ confidence interval may be found by dividing by the population average, μ :

$$\lambda^{(\beta)} = 2b^{(\beta)} \frac{\sigma}{\mu\sqrt{n}} \quad (\text{D.9})$$

$\lambda^{(\beta)}$ is given the name, “ β -level precision”, in this paper (see Section 4.1).

Appendix E. Incorporating a multiplier into the Utility Model: the Extended Utility Model

To allow for variations between the responses of individuals with similar levels of wealth, an additional multiplier, R , may be applied to the VPF, V , derived in Section 3.1, so as to produce a modified VPF, V_m :

$$V_m = RV \quad (\text{E.1})$$

(cf. Eq. (12)). The Extended Utility Model is designed to produce the same expected value as the standard Utility Model, viz.

$$E(V_m) = E(V) \quad (\text{E.2})$$

The inclusion of the multiplier into the model follows the general mathematical development of Section 3.2. Thus the multiplier, R , will have the form given in Eq. (13), the probability density function, $f_Q(q)$, for Q will be as given in Eq. (14), and the expected value of the multiplier will correspond to Eq. (15).

As R and V are modelled as independent, the expected value of their product is simply the product of their expected values:

$$E(V_m) = E(R)E(V) \quad (\text{E.3})$$

Substituting Eq. (E.2) into Eq. (E.3) means that

$$E(R) = 1 \quad (\text{E.4})$$

Combining Eqs. (E.4) and (15) gives the relationship between the interval length, s , and the deterministic component, r_0 , of the multiplier:

$$s = 2(1 - r_0) \quad (\text{E.5})$$

The constraint that $s \geq 0$ means that the smallest value of s , namely zero, occurs when

$$r_0 = 1 = r_{0\max} = E(R) \quad (\text{E.6})$$

Conversely,

$$s = s_{\max} = 2 \text{ when } r_0 = 0 = r_{0\min} \quad (\text{E.7})$$

The variance of R is given by Eq. (20): $\text{var}(R) = s^2/12$.

The variance of V_m is the variance of the product of the two independent random variables, R and V , which may be written [52]:

$$\text{var}(V_m) = \text{var}(RV) = \text{var}(R)\text{var}(V) + (E(V))^2\text{var}(R) + (E(R))^2\text{var}(V) \quad (\text{E.8})$$

(cf. Eq. (22)).

For the Extended Utility Model, the minimum sample size, $n_{V_m\min}$, under the 12% rule may be found by replacing V by V_m in Eq. (30) and using inequality (29) so that

$$n_{V_m\min} = \frac{\text{var}(V_m)}{(E(V_m))^2} \left(\frac{3.92}{0.12}\right)^2 \quad (\text{E.9})$$

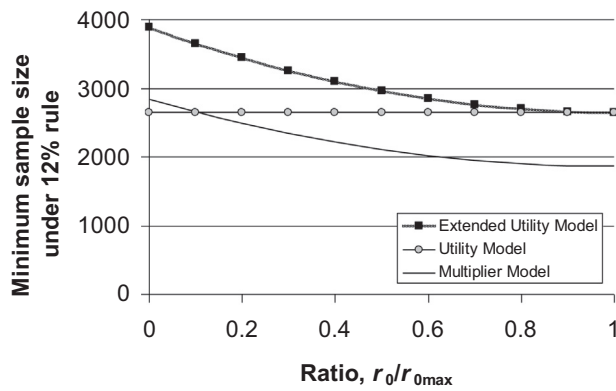


Fig. 12. Minimum sample size, n_{vmin} , under the 12% rule for the Extended Utility Model for VPF survey, for the Utility Model and for the Multiplier Model.

Eq. (E.9) may be evaluated for n_{vmin} as the deterministic component, r_0 , varies between $r_0 = r_{0min} = 0$ and $r_0 = r_{0max} = 1$. It is found that

$$2648 \leq n_{vmin} \leq 3887 \quad (E.10)$$

and that the middle of the range for r_0 gives a value of n_{vmin} that is just below 3000.

Fig. 12 plots the minimum sample size for the Extended Utility Model, n_{vmin} , against the ratio r_0/r_{0max} . This allows the results for the Utility Model and the Multiplier Model to be presented on the same graph. (In the case of the Extended Utility Model, $r_{0max} = 1$ and so $r_0/r_{0max} = r_0$).

References

- [1] L. Finkelstein, Widely, strongly and weakly defined measurement, *Measurement* 34 (2003) 39–48.
- [2] L. Finkelstein, Problems of measurement in soft systems, *Measurement* 38 (2005) 267–274.
- [3] E.L. Glaeser, 2011. Sciences still young, *Harvard Magazine*, September – October, <https://harvardmagazine.com/2011/09/sciences-still-young>.
- [4] J.K. Galbraith, A history of economics. The past as the present, Hamish Hamilton, London, 1987, republished by Penguin Books, London, 1989, pp. 246–247.
- [5] Encyclopaedia Britannica, 1998, revised 2001. Simon Kuznets, American economist and statistician <https://www.britannica.com/biography/Simon-Kuznets>.
- [6] Cambridge Assessment, 2008. How have school exams changed over the past 150 years?, 14 February. <http://www.cambridgeassessment.org.uk/news/how-have-school-exams-changed-over-the-past-150-years/>.
- [7] L. Mari, M. Wilson, An introduction to the Rasch measurement approach for metrologists, *Measurement* 51 (2014) 315–327.
- [8] A. Maul, L. Mari, M. Wilson, Intersubjectivity of measurement across the sciences, *Measurement* 131 (2019) 764–770.
- [9] G.B. Rossi, B. Berglund, Measurement involving human perception and interpretation, *Measurement* 44 (2011) 815–822.
- [10] L. Finkelstein, Measurement: fundamental principles, in: L. Finkelstein, K.T.V. Grattan (Eds.), *Concise Encyclopaedia of Measurement and Instrumentation*, Pergamon Press, Oxford, 1994, pp. 201–205.
- [11] G.B. Rossi, F. Crenna, A formal theory of the measurement system, *Measurement* 116 (2018) 644–651.
- [12] S.S. Stevens, On the theory of scales of measurement, *Science* 103 (2648) (1946) 677–680.
- [13] J. von Neumann, O. Morgenstern, 2007. *Theory of Games and Economic Behavior*, 60th anniversary ed., Princeton University Press, Princeton and Oxford (First edition published 1944), Chapter 1, Sections 3.2 to 3.4.
- [14] P.J. Thomas, An absolute scale for measuring the utility of money, in: Proc. 13th IMEKO TC1-TC7 Joint Symposium Without Measurement no Science, without Science no Measurement, September 1–3, City University, London, UK, IOP Publishing, *Journal of Physics: Conference Series* 238 (2010) 012039 doi: 10.1088/1742-6596/238/1/012039. <http://iopscience.iop.org/article/10.1088/1742-6596/238/1/012039/pdf>.
- [15] P.J. Thomas, Measuring risk-aversion: the challenge, *Measurement* 79 (2016) 285–301, <https://doi.org/10.1016/j.measurement.2015.07.056>.
- [16] J.W. Pratt, Risk aversion in the small and in the large, *Econometrica* 32 (1–2) (1964) 122–136.
- [17] A.B. Atkinson, On the measurement of inequality, *J. Econ. Theor.* 2 (1970) 244–263.
- [18] D.R. Schley, E. Peters, Assessing “economic value”: symbolic-number mappings predict risky and riskless valuations, *Psychol. Sci.* 25 (3) (2014) 753–761.
- [19] P. Thomas, A. Chrystal, Explaining the “buy one get one free” promotion: the golden ratio as a marketing tool, *Am. J. Ind. Bus. Manage.* 3 (8) (2013) 655–673.
- [20] P. Thomas, I. Waddington, 2017. Validating the J-value safety assessment tool against pan-national data, *Process Safety and Environmental Protection*, Vol. 112A, 179–197, November. Available at: https://ac.els-cdn.com/S0957582017302896/1-s2.0-S0957582017302896-main.pdf?_tid=e4b52f78-d2c5-11e7-b19e-00000aacb360&acdnat=1511713484_c1408cbb7d734eb48189c8bec89fcd68.
- [21] P. Thomas, Corroboration of the J-value model for life-expectancy growth in industrialised countries Available at: *Nanotechnol. Percep.* 13 (2017) 31–44. <http://www.jvalue.co.uk/papers/2017-CorroborationOfJvalue.pdf>.
- [22] P.J. Thomas, D.W. Stupples, M.A. Alghaffar, The extent of regulatory consensus on health and safety expenditure. Part 1: Development of the J-value technique and evaluation of regulators’ recommendations, *Trans. IChemE, Part B: Process Saf. Environ. Protect.* 84 (B5) (2006) 329–336.
- [23] P.J. Thomas, R.D. Jones, J.O. Kearns, The trade-offs embodied in J-value analysis, *Process Saf. Environ. Protect.* 88 (3) (2010) 147–167.
- [24] L. Mari, V. Lazzarotti, R. Manzini, Measurement in soft systems: epistemological framework and a case study, *Measurement* 42 (2009) 241–253.
- [25] Sir K.R. Popper, *Logik der Forschung*, translated as *The Logic of Scientific Discovery*, 1959, revised edition 1980, Hutchinson, London, 1934.
- [26] P.J. Thomas, Structural view independence: a criterion for judging the objectivity of economic parameters measured by opinion survey, *Measurement* 47 (2014) 161–177.
- [27] P.J. Thomas, Testing the impartiality of surveys to measure differential risk perception, *Measurement* 60 (2015) 155–173. <http://www.sciencedirect.com/science/article/pii/S026322411400459X>.
- [28] PriceWaterhouseCoopers, 2008. Commentary US Financial Accounting Standards Board Statement FAS 157 – Fair Value Measurement https://www.pwc.com/bm/en/publication/assets/usgaap_08_04.pdf.
- [29] J. Locke, 1690. *Essay Concerning Human Understanding*, London, <http://www.earlymoderntexts.com/assets/pdfs/locke1690book1.pdf>.
- [30] D. Fujiwara, R. Campbell, 2011. Valuation techniques for social cost-benefit analysis https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/209107/greenbook_valuationtechniques.pdf.
- [31] P.J. Thomas, 2018. Pitfalls in survey measurements of economic parameters, in: XXII World Congress of the International Measurement Confederation (IMEKO), Belfast, 3–6 September, J. Phys.: Conf. Ser. 1065 072009. <http://iopscience.iop.org/article/10.1088/1742-6596/1065/7/072009/pdf>.
- [32] P. Thomas, G. Vaughan, All in the balance: assessing schemes to protect humans and the environment, *Nucl. Future* 9 (3) (2013) 41–51.
- [33] T. Carthy, S. Chilton, J. Covey, L. Hopkins, M. Jones-Lee, G. Loomes, N. Pidgeon, A. Spencer, On the contingent valuation of safety and the safety of contingent valuation: Part 2 – The CV/SG “chained” approach, *J. Risk Uncertainty* 17 (3) (1999) 187–213.
- [34] P.J. Thomas, G.J. Vaughan, Testing the validity of the “value of a prevented fatality” (VPF) used to assess UK safety measures, *Process Saf. Environ. Protect.* 94 (2015) 239–261. <http://www.sciencedirect.com/science/article/pii/S0957582014000962>.
- [35] U.S. Department of Transportation, 2016. Guidance on Treatment of the Economic Value of a Statistical Life in U.S. Department of Transportation Analyses – 2016 Adjustment, 8 August. <https://www.transportation.gov/sites/dot.gov/files/docs/2016%20Revised%20Value%20of%20a%20Statistical%20Life%20Guidance.pdf>.
- [36] S. Chilton, J. Covey, M. Jones-Lee, G. Loomes, N. Pidgeon, A. Spencer, Response to Thomas and Vaughan, *Process Saf. Environ. Protect.* 93 (2015) 293–298.
- [37] M. Jones-Lee, G. Loomes, Final response to Thomas and Vaughan, *Process Saf. Environ. Protect.* 94 (2015) 542–544.
- [38] P.J. Thomas, G.J. Vaughan, “Testing the validity of the “value of a prevented fatality” (VPF) used to assess UK safety measures”: reply to the comments of Chilton, Covey, Jones-Lee, Loomes, Pidgeon and Spencer, *Process Saf. Environ. Protect.* 93 (2015) 299–306. <http://www.sciencedirect.com/science/article/pii/S0957582014001797>.
- [39] P.J. Thomas, G.J. Vaughan, Pitfalls in the application of utility functions to the valuation of human life, *Process Saf. Environ. Protect.* 98 (2015) 148–169. <http://www.sciencedirect.com/science/article/pii/S095758201500124X>.
- [40] P. Thomas, I. Waddington, 2017. What is the value of life? A review of the value of a prevented fatality used by regulators and others in the UK, *Nucl. Future*, 13(1), 32–39. Available at: <http://www.bristol.ac.uk/cabot/news/2017/research-review-safety.html> and http://www.wcsim.co.uk/files/8314/8767/9271/What_is_the_Value_of_Life_-_Published_pdf.pdf.
- [41] P. Thomas, G. Vaughan, 2014. Explaining perceived inconsistencies in “Stated Preference”. Valuations of human life, *Am. J. Ind. Bus. Manage.*, 4(9), 442–473. <http://www.scirp.org/journal/PaperInformation.aspx?paperID=49597&#abstract>.
- [42] J. Beattie, J. Covey, P. Dolan, L. Hopkins, M. Jones-Lee, G. Loomes, N. Pidgeon, A. Robinson, A. Spencer, On the contingent valuation of safety and the safety of contingent valuation: Part 1 – Caveat investigator, *J. Risk Uncertain.* 17 (1998) 5–25.
- [43] ONS, 2018. *Wealth in Great Britain Wave 5: 2014 to 2016*, 1 February. <https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/bulletins/wealthingreatbritainwave5/2014to2016>.
- [44] BBC, 2018. Rich List 2018: Jim Ratcliffe is UK’s richest man, May 13, <https://www.bbc.co.uk/news/uk-44096507>.

- [45] ONS, 2017. Families and Households: 2017, November 8, <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/families/bulletins/familiesandhouseholds/2017>.
- [46] W.K. Viscusi, W.N. Evans, Utility functions that depend on health status: estimates and economic implications, *Am. Econ. Rev.* 80 (3) (1990) 353–374.
- [47] C.R. Sunstein, 2003. Lives, Life-Years, and Willingness to Pay, University of Chicago John M. Olin Law & Economics Working Paper No. 191, https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=1310&context=law_and_economics (accessed August 2018)
- [48] D. Kahneman, A. Tversky, Prospect theory: an analysis of decision under risk, *Econometrica* 47 (1979) 263–291.
- [49] D. Kahneman, *Thinking, Fast and Slow*, Penguin Books, London, U.K., 2011.
- [50] H.M. Treasury, 2018. The Green Book: Central Government Guidance on Appraisal and Evaluation, Appendix A6, Discounting, paragraph A6.12, available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/685903/The_Green_Book.pdf.
- [51] H.M. Treasury, 2011. The Green Book: Central Government Guidance on Appraisal and Evaluation, 2003 Edition updated July 2011, Annex 5, Box 5.1.
- [52] C.L. Chiang, *Introduction to Stochastic Processes in Biostatistics*, John Wiley, New York, USA, 1968.
- [53] R.J. Larsen, M.L. Marx, *An introduction to Mathematical Statistics and Its Applications*, 3rd Ed., Prentice Hall, New Jersey, USA, 2001, pp. 265–273.
- [54] Electoral Commission, 2016, EU Referendum Results, available at: <https://www.electoralcommission.org.uk/find-information-by-subject/elections-and-referendums/past-elections-and-referendums/eu-referendum/electorate-and-count-information> Accessed 2 March 2019.